

CANADIAN
JOURNAL OF MATHEMATICS*Journal Canadien de Mathématiques*

VOL. IX · NO. 1

1957

<i>New characterizations of polyhedral cones</i>	H. Mirkil	1
<i>Some global theorems on hypersurfaces</i>	Chuan-Chih Hsiung	5
<i>A note on the Mathieu groups</i>	Lowell J. Paige	15
<i>Relative cohomology</i>	D. G. Higman	19
<i>Some remarks on Noetherian rings</i>	Michio Yoshida	35
<i>Spaces of dimension zero</i>	Bernhard Banaschewski	38
<i>Matrices with elements in a Boolean ring</i>	A. T. Butson	47
<i>Characteristic polynomials</i>	Hans Schneider	60
<i>Some theorems about $p_r(n)$</i>	Morris Newman	68
<i>Classes of positive definite unimodular circulants</i>	Morris Newman and Olga Taussky	71
<i>Simultaneous pairs of linear and quadratic equations in a Galois field</i>	Eckford Cohen	74
<i>The set of all generalized limits of bounded sequences</i>	Meyer Jerison	79
<i>A note on asymptotic series</i>	H. F. Davis	90
<i>On Ward's Perron Stieltjes integral</i>	Ralph Henstock	96
<i>A generalization of the Cauchy principal value</i>	Charles Fox	110
<i>On the zeros of the Fresnel integrals</i>	Erwin Kreyszig	118
<i>A minimum-maximum problem for differential expressions</i>	D. S. Carter	132
<i>A mixed problem for normal hyperbolic linear partial differential equations of second order</i>	G. F. D. Duff	141

Published for

THE CANADIAN MATHEMATICAL CONGRESS

by the

University of Toronto Press

EDITORIAL BOARD

H. S. M. Coxeter, A. Gauthier, R. D. James, R. L. Jeffery,
G. de B. Robinson, H. Zassenhaus

with the co-operation of

H. Behnke, R. Brauer, D. B. DeLury, G. F. D. Duff, I. Halperin,
W. K. Hayman, J. Leray, S. MacLane, P. Scherk, B. Segre,
J. L. Synge, W. J. Webber

The chief languages of the *Journal* are English and French.

Manuscripts for publication in the *Journal* should be sent to the *Editor-in-Chief*, H. S. M. Coxeter, University of Toronto. Everything possible should be done to lighten the task of the reader; the notation and reference system should be carefully thought out. Every paper should contain an introduction summarizing the results as far as possible in such a way as to be understood by the non-expert.

All other correspondence should be addressed to the *Managing Editor*, G. de B. Robinson, University of Toronto.

The *Journal* is published quarterly. Subscriptions should be sent to the *Managing Editor*. The price per volume of four numbers is \$8.00. This is reduced to \$4.00 for individual members of recognized Mathematical Societies.

The Canadian Mathematical Congress gratefully acknowledges the assistance of the following towards the cost of publishing this *Journal*:

University of Alberta	Assumption University
University of British Columbia	Carleton College
Dalhousie University	Ecole Polytechnique
Université Laval	Loyola College
University of Manitoba	McGill University
McMaster University	Université de Montréal
Queen's University	Royal Military College
St. Mary's University	University of Toronto
National Research Council of Canada	
and the	
American Mathematical Society	

he
ng
on
er
le

ng

at
rs
of

he
l:

NI

1

who
Not
(x,
the
upp
to s
of i
crit
sub

T
and
mu
eve
is (
ting
call
We
 $f(x)$
gen
Th
tha
con

(1)

(2)

(3)

Th

for

2

if i

(3)

and

T

clo

pro

R

NEW CHARACTERIZATIONS OF POLYHEDRAL CONES

H. MIRKIL

1. Introduction. A pyramid clearly has all its projections closed, even when the line segments from vertex to base are extended to infinite half-lines. Not so a circular cone. For if the cone is on its side and supported by the (x, y) plane in such a way that its infinite half-line of support coincides with the positive x axis, then its horizontal projection on the (y, z) plane is the open upper half-plane $y > 0$, together with the single point $(0, 0)$. It is our purpose to show that the pyramid behaves better under projection precisely because of its *polyhedral* nature. And this principle can be reinterpreted to give a criterion for the positive extendibility of positive functionals defined on a subspace of a partially ordered vector space.

Throughout our discussion E will be a real finite-dimensional vector space and E' its dual space. A subset P of E stable under vector addition and multiplication by non-negative scalars is called a *convex cone*. In particular, every linear subspace is a convex cone. The smallest subspace containing P is $(-P) + P$, and the largest subspace contained in P is $(-P) \cap P$. Omitting parentheses, we shall write $-P + P$ and $-P \cap P$. It is customary to call $\dim(-P + P)$ the *dimension* of P , and $\dim(-P \cap P)$ the *lineality* of P . We define the *polar* P° of P to be the set of all functionals $f \in E'$ such that $f(x) \geq 0$ for all $x \in P$. P° is a closed convex cone, and in fact is the most general such cone, since the double polar $P^{\circ\circ}$ coincides with the closure of P . This fact authorizes us to use the notation $P^{\circ\circ}$ for the closure of P (provided that P is a convex cone). The elementary duality theory of closed convex cones can be summed up as follows:

- (1) *Galois connection*: $(P + Q)^\circ = P^\circ \cap Q^\circ$ and $(P \cap Q)^\circ = (P^\circ + Q^\circ)^{\circ\circ}$.
- (2) $\dim(-P + P) + \dim(-P^\circ + P^\circ) = \dim E$.
- (3) If $P^\circ = E'$, then $P = E$.

The above statements for closed convex cones P and Q are easily rephrased for arbitrary convex cones. Proofs can be found in Fenchel (2).

2. Cones with all projections closed. We shall call a cone *polyhedral* if it is the intersection of finitely many closed half-spaces. A theorem of Weyl (3) states that such a cone is also the convex hull of finitely many half-lines, and conversely. Thus P and P° are polyhedral together.

THEOREM. *If a closed convex cone P has all its 2-dimensional projections closed, then P is polyhedral. Conversely, a closed convex polyhedral cone P has its projections (of all dimensions) closed.*

Received February 22, 1956.

Proof. We first dispose of the converse. Let P be polyhedral and let T be a projector. Then P is precisely all non-negative combinations of finitely many fixed vectors x_1, \dots, x_m and TP is all non-negative combinations of Tx_1, \dots, Tx_m . Hence TP is not only closed but even polyhedral.

Now assume that P and all its 2-dimensional projections are closed and make an induction on the dimension n of $-P + P$, which we can without loss of generality take equal to the dimension of the whole vector space E . Since there are only 6 linearly inequivalent closed convex cones of dimension ≤ 2 , and since all of these are polyhedral, we can begin with $n \geq 3$.

First suppose that P contains some line L , or equivalently that $\dim(-P \cap P) > 0$. If T projects P parallel to L into some hyperplane (L the null space of the projector T), then TP has lower dimension than P and has all its 2-dimensional projections closed; hence by induction TP is the smallest cone containing the images Tx_1, \dots, Tx_s of some $x_1, \dots, x_s \in P$. Let x be any non-zero vector in L . Because $P = P + L$, then P is the complete inverse image of TP . Hence P is the smallest cone containing $x, -x, x_1, \dots, x_s$.

Suppose now $\dim(-P \cap P) = 0$. The convex cone $-P^0 + P^0$ is dense in E' , hence is all of E' , and P^0 contains n linearly independent functionals whose sum f is strictly positive at all non-zero points of P . Let H be the affine hyperplane $\{x: f(x) = 1\}$. P is the smallest cone containing $H \cap P$, and this compact convex set is the convex hull of its extreme points. We shall show that these are finite in number by showing that each point $x \in H \cap P$ has a neighbourhood containing no extreme point other than possibly x itself.

Let H_0 be the linear hyperplane through the origin parallel to H and let L be the line through x . Let T project the whole space E onto H_0 along L . The cone TP is polyhedral by induction, hence consists of all non-negative combinations of some $y_1, \dots, y_r \in H_0$. We can without loss of generality take these as images under T of

$$y_1 + x, \dots, y_r + x \in H \cap P.$$

(For if they are originally the images of $y_1 + c_1x, \dots, y_r + c_rx \in P$, then we can replace y_1, \dots, y_r by $c_1^{-1}y_1, \dots, c_r^{-1}y_r$.)

Furthermore, near 0, TP coincides with the convex hull of 0, y_1, \dots, y_r . We assert that near x the set $H \cap P$ coincides with the convex hull of x, x_1, \dots, x_r . For clearly any $x' \in H \cap P$ near x has its image Tx' near 0, hence $Tx' = c_1y_1 + \dots + c_r y_r$, with all c 's non-negative and $0 \leq c_1 + \dots + c_r < 1$. And since the restriction of T to H is a one-one affine mapping of H onto H_0 , it follows that x' is the convex combination

$$(1 - c_1 - \dots - c_r)x + c_1x_1 + \dots + c_rx_r.$$

Now $H \cap P$ is seen to be a convex polyhedron near x , and near x the only possible extreme point of $H \cap P$ is x itself. Thus $H \cap P$ coincides near x with the convex hull of x, x_1, \dots, x_r .

An application of the Heine-Borel theorem completes the proof. We cover the compact convex set $H \cap P$ with finitely many open sets, each containing at most one extreme point of $H \cap P$. The whole set $H \cap P$ is then the convex hull of finitely many points, P is the convex hull of finitely many half-lines, and the theorem is proved.

3. Positive extendibility of positive functionals. The main practical interest of our theorem lies in Corollary 1 below. De Leeuw (1) uses it, and a good deal more, in proving a convexity theorem for polynomials in several complex variables. And in fact it was a question of his about positive extendibility of functionals that led us to conjecture the theorem of the present paper.

Given a convex cone P that contains no line, we can make E an *ordered vector space* by defining $x \succ y$ to mean $x - y \in P$. Conversely the vectors $\succ 0$ in an ordered vector space form a convex cone P that contains no line. When such a cone is closed it is called an *order cone*. And it is easy to prove that each of the following properties characterizes order cones among all closed convex cones P :

- (1) P possesses extreme half-lines (though certain authors define extreme half-lines in such a way that this particular characterization has trivial exceptions).
- (2) P is the convex hull of its extreme half-lines.
- (3) P lies strictly to one side of some hyperplane.
- (4) P consists of all half-lines through some compact convex set that does not contain the origin.

Our purpose in this section, however, is to characterize polyhedral order cones among all order cones. The simplest kind of polyhedral ordering of a vector space is the coordinatewise ordering relative to a basis, a vector being non-negative if and only if all its coordinates are non-negative. A natural example of non-polyhedral (and non-lattice) ordering is given by the cone of positive-definite matrices in the real n^2 -dimensional space of n -by- n complex hermitian matrices.

COROLLARY 1. *Let E be a polyhedrally ordered vector space, and let F be a subspace with the induced ordering. Then every positive functional f on F can be extended to a positive functional on E . Conversely, let E be a vector space ordered by a closed cone in such a way that the above positive extension property holds for all subspaces F and positive functionals f on F . Then the ordering of E is polyhedral.*

Proof. Let T_F be the natural mapping of E' onto F' , with null space F° . The condition for positive extendibility of functionals can be written

$$T_F(P^\circ) = T_F((P \cap F)^\circ) = T_F((P^\circ + F^\circ)^{\circ\circ})$$

or equivalently

$$P^\circ + F^\circ = (P^\circ + F^\circ)^{\circ\circ} + F^\circ = (P^\circ + F^\circ)^{\circ\circ}.$$

Thus the condition asserts that all projections of P° are closed, and we know by our theorem that this happens exactly when P° is polyhedral. But P and P° are polyhedral together.

COROLLARY 2. *Let F be a subspace of E , let P be a polyhedral cone in E containing no line, and let K_F be a half-space of F containing $F \cap P$. Then $K_F = F \cap K_E$, for some half-space of E containing P . Conversely, let P be a closed convex cone containing no line, and suppose that some half-space K_E of E like the above can be found for every choice of F and K_F . Then P is polyhedral.*

Proof. We have simply restated Corollary 1 in a geometric language that avoids explicit mention of functionals.

REFERENCES

1. K. de Leeuw, *A type of convexity in the space of n complex variables*, Trans. Amer. Math. Soc. (to appear).
2. W. Fenchel, *Convex cones, sets, and functions*, Princeton lecture notes (mimeographed) 1953.
3. H. Weyl, *Elementäre Theorie der konvexen Polyeder*, Comm. Math. Helv. 7 (1935), 290.

*Dartmouth College
Hanover, N.H.*

SOME GLOBAL THEOREMS ON HYPERSURFACES

CHUAN-CHIH HSIUNG

1. Introduction. The purpose of this paper is to establish the following theorems, which were obtained by Hopf and Voss in their joint paper (2) for the case where $n = 2$.

THEOREM 1. *Let V^n, V^{*n} be two closed orientable hypersurfaces twice differentiably imbedded in a Euclidean space E^{n+1} of dimension $n + 1 \geq 3$. Suppose that there is a differentiable homeomorphism between the two hypersurfaces V^n, V^{*n} such that the orientations of the two hypersurfaces V^n, V^{*n} are preserved and the line joining every pair of corresponding points P, P^* of the two hypersurfaces V^n, V^{*n} is parallel to a fixed direction R , and such that the two hypersurfaces V^n, V^{*n} have equal first mean curvatures at every pair of the points P, P^* but no cylindrical elements whose generators are parallel to the fixed direction R . Then the two hypersurfaces V^n, V^{*n} can be transformed into each other by a translation.*

A closed hypersurface V^n imbedded in a Euclidean space E^{n+1} of dimension $n + 1 \geq 2$ is said to be *convex in a given direction*, if no line in this direction intersects the hypersurface V^n at more than two points. It is obvious that a closed hypersurface V^n is convex in the usual sense if it is convex in every direction in the space E^{n+1} .

THEOREM 2. *Let a closed orientable hypersurface V^n twice differentiably imbedded in a Euclidean space E^{n+1} of dimension $n + 1 \geq 3$ be convex in a given direction R . If the two first mean curvatures of the hypersurface V^n at every pair of its points of intersection with the lines in the direction R are equal, then the hypersurface V^n has a hyperplane of symmetry perpendicular to the direction R .*

Theorem 2 can easily be deduced from Theorem 1. In fact, let u be a mapping of a hypersurface V^n satisfying the conditions of Theorem 2 onto itself such that the two points of intersection of the hypersurface V^n with any line in the direction R are mapped into each other. In particular, if a line in the direction R is tangent to the hypersurface V^n at a point P , then $uP = P$. Let r be the reflection with respect to an arbitrary hyperplane perpendicular to the direction R , and P any point of the hypersurface V^n . Then the mapping $ruP = P^*$ maps the hypersurface V^n onto the hypersurface $V^{*n} = r(V^n)$ generated by the point P^* , and the two hypersurfaces V^n, V^{*n} satisfy the conditions of Theorem 1 so that $ru = t$ is a translation. Therefore $u = rt$ is a reflection with respect to a hyperplane perpendicular to the direction R , and hence Theorem 2 follows.

Received April 7, 1956.

By noting that a closed hypersurface V^n imbedded in a Euclidean space E^{n+1} of dimension $n + 1 \geq 2$ must be a hypersphere if it has a hyperplane of symmetry perpendicular to every direction in the space E^{n+1} , we arrive readily at the following known result from Theorem 2.

COROLLARY. *A closed convex hypersurface V^n of constant first mean curvature twice differentially imbedded in a Euclidean space E^{n+1} of dimension $n + 1 \geq 3$ is a hypersphere.*

THEOREM 3. *Let $V^n(V^{*n})$ be an orientable hypersurface with a closed boundary $V^{n-1}(V^{*n-1})$ of dimension $n - 1 \geq 1$ twice differentially imbedded in a Euclidean space E^{n+1} of dimension $n + 1$. Suppose that there is a differentiable homeomorphism between the two hypersurfaces V^n, V^{*n} with the same properties as those of the homeomorphism in Theorem 1.*

(i) *If the two boundaries V^{n-1}, V^{*n-1} are coincident, then the two hypersurfaces V^n, V^{*n} are coincident.*

(ii) *If the two normals of the two hypersurfaces V^n, V^{*n} at every pair of corresponding points, under the given homeomorphism, of the two boundaries V^{n-1}, V^{*n-1} are parallel, then the two hypersurfaces V^n, V^{*n} are transformed into each other by a translation.*

2. Preliminaries¹. In a Euclidean space E^{n+1} of dimension $n + 1 \geq 3$, let us consider a fixed orthogonal frame $OI_1 \dots I_{n+1}$ with a point O as the origin. With respect to this orthogonal frame we define the vector product of n vectors A_1, \dots, A_n in the space E^{n+1} to be the vector A_{n+1} , denoted by $A_1 \times \dots \times A_n$, satisfying the following conditions:

(a) the vector A_{n+1} is normal to the n -dimensional subspace of E^{n+1} determined by the vectors A_1, \dots, A_n ,

(b) the magnitude of the vector A_{n+1} is equal to the volume of the parallelepiped whose edges are the vectors A_1, \dots, A_n ,

(c) the two frames $OA_1 \dots A_n A_{n+1}$ and $OI_1 \dots I_{n+1}$ have the same orientation.

Let σ be a permutation on the n numbers $1, \dots, n$, then

$$(2.1) \quad A_{\sigma(1)} \times \dots \times A_{\sigma(n)} = (\text{sgn } \sigma) A_1 \times \dots \times A_n,$$

where $\text{sgn } \sigma$ is $+1$ or -1 according as the permutation σ is even or odd. Let i_1, \dots, i_{n+1} be the unit vectors from the origin O in the directions of the vectors I_1, \dots, I_{n+1} and let A_α^j ($j = 1, \dots, n + 1$) be the components² of the vector A_α ($\alpha = 1, \dots, n$) with respect to the frame $OI_1 \dots I_{n+1}$, then the scalar product of any two vectors A_α and A_β and the vector product of n vectors A_1, \dots, A_n are, respectively,

¹For this section see, for instance, (3, pp. 287-289).

²Throughout this paper all Latin indices take the values 1 to $n + 1$ and Greek indices the values 1 to n unless stated otherwise. We shall also follow the convention that repeated indices imply summation.

$$(2.2) \quad A_\alpha \cdot A_\beta = \sum_{i=1}^{n+1} A_\alpha^i A_\beta^i,$$

$$(2.3) \quad A_1 \times \dots \times A_n = (-1)^n \begin{vmatrix} i_1 & i_2 & \dots & i_{n+1} \\ A_1^1 & A_1^2 & \dots & A_1^{n+1} \\ \dots & \dots & \dots & \dots \\ A_n^1 & A_n^2 & \dots & A_n^{n+1} \end{vmatrix}.$$

If A_α^i are differentiable functions of n variables x^1, \dots, x^n , then by equation (2.3) and the differentiation of determinants

$$(2.4) \quad \frac{\partial}{\partial x^\alpha} (A_1 \times \dots \times A_n) = \sum_{\beta=1}^n \left(A_1 \times \dots \times A_{\beta-1} \times \frac{\partial A_\beta}{\partial x^\alpha} \times A_{\beta+1} \times \dots \times A_n \right).$$

Now we consider a hypersurface V^n twice differentially imbedded in the space E^{n+1} with a closed boundary V^{n-1} of dimension $n-1$. Let (y^1, \dots, y^{n+1}) be the coordinates of a point P in the space E^{n+1} with respect to the orthogonal frame $OI_1 \dots I_{n+1}$. Then the hypersurface V^n can be given by the parametric equations

$$(2.5) \quad y^i = f^i(x^1, \dots, x^n) \quad (i = 1, \dots, n+1),$$

or the vector equation

$$(2.6) \quad Y = F(x^1, \dots, x^n),$$

where y^i and f^i are respectively the components of the two vectors Y and F , the parameters x^1, \dots, x^n take values in a simply connected domain D of the n -dimensional real number space, $f^i(x^1, \dots, x^n)$ are twice differentiable and the Jacobian matrix $||\partial y^i / \partial x^\alpha||$ is of rank n at all points of the domain D . If we denote the vector $\partial Y / \partial x^\alpha$ by Y_α ($\alpha = 1, \dots, n$), then the first fundamental form of the hypersurface V^n at the point P is

$$(2.7) \quad ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta,$$

where

$$(2.8) \quad g_{\alpha\beta} = Y_\alpha \cdot Y_\beta,$$

and the matrix $||g_{\alpha\beta}||$ is positive definite so that the determinant $g = |g_{\alpha\beta}| > 0$.

Let N be the unit normal vector of the hypersurface V^n at the point P , and N_α the vector $\partial N / \partial x^\alpha$, then

$$(2.9) \quad N_\alpha = -b_{\alpha\beta} g^{\beta\gamma} Y_\gamma,$$

where

$$(2.10) \quad b_{\alpha\beta} = b_{\beta\alpha} = -N_\alpha \cdot Y_\beta$$

are the coefficients of the second fundamental form of the hypersurface V^n at the point P , and $g^{\beta\gamma}$ denotes the cofactor of $g_{\beta\gamma}$ in g divided by g so that

$$(2.11) \quad g^{\alpha\beta} g_{\beta\gamma} = \delta_{\gamma}^{\alpha}$$

(the Kronecker deltas). The n principal curvatures $\kappa_1, \dots, \kappa_n$ of the hypersurface V^n at the point P are the roots of the determinant equation

$$(2.12) \quad |b_{\alpha\beta} - \kappa g_{\alpha\beta}| = 0,$$

from which follows immediately the first mean curvature of the hypersurface V^n at the point P :

$$(2.13) \quad M_1 = \frac{1}{n} \sum_{\alpha=1}^n \kappa_{\alpha} = \frac{1}{n} b_{\alpha\beta} g^{\alpha\beta}.$$

The area element of the hypersurface V^n at the point P is

$$(2.14) \quad dA = g^{\frac{1}{2}} dx^1 \wedge \dots \wedge dx^n,$$

where the operator d is the exterior differentiation, and the wedge denotes the exterior multiplication. Now we choose the direction of the unit normal vector N in such a way that the two frames $PY_1 \dots Y_n N$ and $OI_1 \dots I_{n+1}$ have the same orientation. Then from equations (2.3) and (2.14) it follows that

$$(2.15) \quad g^{\frac{1}{2}} N = Y_1 \times \dots \times Y_n,$$

$$(2.16) \quad |Y_1, \dots, Y_n, N| = g^{\frac{1}{2}},$$

where the left side of equation (2.16) is a determinant indicated by writing only a typical row.

3. An integral formula. Let V^n be an orientable hypersurface with a closed boundary V^{n-1} of dimension $n-1 \geq 1$ twice differentially imbedded in a Euclidean space E^{n+1} of dimension $n+1$, and suppose that the hypersurface V^n is given by the vector equation (2.6). Let I be the unit vector in a fixed direction R in the space E^{n+1} , and w a twice differentiable function over the hypersurface V^n . Then §2 can be applied to the hypersurface V^n , and we shall use the same symbols with a star for the corresponding quantities for the hypersurface V^{*n} defined by the vector equation

$$(3.1) \quad Y^* = Y + W,$$

where

$$(3.2) \quad W = wI.$$

Let Ω^{α} ($\alpha = 1, \dots, n$) be n vectors in the space E^{n+1} , and suppose that the components of each vector Ω^{α} with respect to the orthogonal frame $OI_1 \dots I_{n+1}$ are differentiable functions of the n variables x^1, \dots, x^n . In order to derive an integral formula for the two hypersurfaces V^n , V^{*n} we use the vector product of vectors and the exterior multiplication of differentials to define the vector

$$(3.3) \quad \Omega^1 \otimes \dots \otimes \Omega^{n-1} \otimes d\Omega^n \otimes \dots \otimes d\Omega^n \\ = (\Omega^1 \times \dots \times \Omega^{n-1} \times \Omega_{\theta_{\alpha}}^{\alpha} \times \dots \times \Omega_{\theta_n}^n) dx^{\theta_1} \wedge \dots \wedge dx^{\theta_n}$$

for $\alpha = 1, \dots, n$, where

$$\Omega_{\alpha}^n = \partial\Omega^n / \partial x^{\alpha}.$$

It is obvious that the vector (3.3) is independent of the order of the vectors $d\Omega^1, \dots, d\Omega^n$. Thus from equations (2.9), (2.13), (2.14), (2.15) we obtain

$$(3.4) \quad dY \otimes \dots \otimes dY = n! (Y_1 \times \dots \times Y_n) dx^1 \wedge \dots \wedge dx^n = n! N dA,$$

$$(3.5) \quad dY \otimes \dots \otimes dY \otimes dN$$

$$= (n-1)! \left(\sum_{\alpha=1}^n Y_1 \times \dots \times Y_{\alpha-1} \times N_{\alpha} \times Y_{\alpha+1} \times \dots \times Y_n \right) dx^1 \wedge \dots \wedge dx^n \\ = -n! M_1 N dA.$$

Making use of equations (3.1), (3.2), (3.4) and its analogue for the hypersurface $V^{\alpha n}$, and noting that

$$dW \otimes \dots \otimes dW \otimes dY \otimes \dots \otimes dY = 0, \\ dW \otimes \dots \otimes dW \otimes dY^* \otimes \dots \otimes dY^* = 0$$

(α factors) (n-α factors) (α factors) (n-α factors)

for $\alpha \geq 2$ and

$$|W, Y_1, \dots, Y_n| = |W, Y_1^*, \dots, Y_n^*|,$$

we are easily led to

$$(3.6) \quad (n-1)! (N^* dA^* - N dA) = dW \otimes dY \otimes \dots \otimes dY \\ = dW \otimes dY^* \otimes \dots \otimes dY^*,$$

$$(3.7) \quad W \cdot N dA = W \cdot N^* dA^*,$$

$$(3.8) \quad |W, N^*, Y_1^*, \dots, Y_{\alpha-1}^*, Y_{\alpha+1}^*, \dots, Y_n^*| \\ = |W, N^*, Y_1, \dots, Y_{\alpha-1}, Y_{\alpha+1}, \dots, Y_n| \quad (\alpha = 1, \dots, n).$$

From equations (2.3), (3.3), (3.5), (3.6) it follows immediately that

$$(3.9) \quad W \cdot (N \otimes dY \otimes \dots \otimes dY) \\ = (-1)^n (n-1)! \sum_{\alpha=1}^n |W, N, Y_1, \dots, Y_{\alpha-1}, Y_{\alpha+1}, \dots, Y_n| \\ dx^1 \wedge \dots \wedge dx^{\alpha-1} \wedge dx^{\alpha+1} \wedge \dots \wedge dx^n,$$

$$(3.10) \quad d[W \cdot (N \otimes dY \otimes \dots \otimes dY)] \\ = -N \cdot (dW \otimes dY \otimes \dots \otimes dY) + W \cdot (dN \otimes dY \otimes \dots \otimes dY) \\ = -n! M_1 W \cdot N dA - (n-1)! (N \cdot N^* dA^* - dA).$$

Similarly, in consequence of equations (3.6), (3.7), (3.8) and those analogous to equations (3.5), (3.9) by changing the vectors Y, N to the vectors Y^*, N^* respectively, we obtain

$$\begin{aligned}
 (3.11) \quad d[W \cdot (N^* \otimes dY \otimes \dots \otimes dY)] &= d[W \cdot (N^* \otimes dY^* \otimes \dots \otimes dY^*)] \\
 &= -N^* \cdot (dW \otimes dY^* \otimes \dots \otimes dY^*) + W \cdot (dN^* \otimes dY^* \otimes \dots \otimes dY^*) \\
 &= -n! M_1^* W \cdot N dA - (n-1)! (dA^* - N^* \cdot N dA).
 \end{aligned}$$

Thus, from equations (3.9), (3.10), (3.11),

$$\begin{aligned}
 (3.12) \quad d \sum_{\alpha=1}^n |W, N - N^*, Y_1, \dots, Y_{\alpha-1}, Y_{\alpha+1}, \dots, Y_n| \\
 dx^1 \wedge \dots \wedge dx^{\alpha-1} \wedge dx^{\alpha+1} \wedge \dots \wedge dx^n \\
 = \frac{(-1)^n}{(n-1)!} d[W \cdot (N \otimes dY \otimes \dots \otimes dY) - W \cdot (N^* \otimes dY \otimes \dots \otimes dY)] \\
 = (-1)^n [n(M_1^* - M_1) W \cdot N dA + (1 - N \cdot N^*) (dA + dA^*)].
 \end{aligned}$$

Integrating equation (3.12) over the hypersurface V^n and applying the Stokes' theorem to the left side of the equation, we then arrive at the integral formula

$$\begin{aligned}
 (3.13) \quad \int_{V^{n-1}} \sum_{\alpha=1}^n |W, N - N^*, Y_1, \dots, Y_{\alpha-1}, Y_{\alpha+1}, \dots, Y_n| \\
 dx^1 \wedge \dots \wedge dx^{\alpha-1} \wedge dx^{\alpha+1} \wedge \dots \wedge dx^n \\
 = (-1)^n \int_{V^n} [n(M_1^* - M_1) W \cdot N dA + (1 - N \cdot N^*) (dA + dA^*)].
 \end{aligned}$$

In particular, when the hypersurface V^n is closed and orientable, the integral on the left side of equation (3.13) vanishes and hence

$$(3.14) \quad n \int_{V^n} (M_1^* - M_1) W \cdot N dA + \int_{V^n} (1 - N \cdot N^*) (dA + dA^*) = 0.$$

4. Proof of Theorems 1 and 3. It is easily seen that we can apply the results in §3 to two hypersurfaces V^n, V^{*n} satisfying the assumptions of Theorem 1. Since $M_1^* = M_1$ at every pair of corresponding points of the two hypersurfaces V^n, V^{*n} , the formula (3.14) becomes

$$(4.1) \quad \int_{V^n} (1 - N \cdot N^*) (dA + dA^*) = 0.$$

But $dA > 0$, $dA^* > 0$ and $1 - N \cdot N^* \geq 0$ due to the fact that N and N^* are unit vectors. Thus the integrand of equation (4.1) is non-negative, and therefore equation (4.1) holds when and only when $1 - N \cdot N^* = 0$, which implies that

$$(4.2) \quad N^* = N.$$

Now in the space E^{n+1} we choose the orthogonal frame $OI_1 \dots I_{n+1}$, with respect to which a point in the space E^{n+1} has coordinates y^1, \dots, y^{n+1} , in such a way that the unit vector I_{n+1} is the fixed unit vector I . Since the hypersurface V^n has no cylindrical elements whose generators are parallel to the fixed vector I , the closed set M of all points of the hypersurface V^n , at each

of which the y^{n+1} -component of the unit normal vector N of the hypersurface V^n is zero, has no inner points and therefore the open set $V^n - M$ is everywhere dense over V^n . Thus, in neighborhoods of any point of the set $V^n - M$ and its corresponding point on the hypersurface V^{*n} , y^1, \dots, y^n are regular parameters of the two hypersurfaces V^n, V^{*n} so that the hypersurfaces V^n, V^{*n} can be represented respectively by the equations

$$(4.3) \quad \begin{aligned} y^{n+1} &= y^{n+1}(y^1, \dots, y^n), \\ y^{*n+1} &= y^{*n+1}(y^1, \dots, y^n) = y^{n+1}(y^1, \dots, y^n) + w(y^1, \dots, y^n). \end{aligned}$$

By means of equations (2.15), (4.3) we obtain the unit normal vectors N, N^* of the hypersurfaces V^n, V^{*n} :

$$(4.4) \quad N = -g^{-1} \left(\sum_{\alpha=1}^n \frac{\partial y^{n+1}}{\partial y^\alpha} i_\alpha - i_{n+1} \right), \quad N^* = -g^{*-1} \left(\sum_{\alpha=1}^n \frac{\partial y^{*n+1}}{\partial y^\alpha} i_\alpha - i_{n+1} \right),$$

from which and equations (4.2), (4.3) it follows immediately that in a neighborhood of any point of the set $V^n - M$,

$$\partial y^{*n+1} / \partial y^\alpha = \partial y^{n+1} / \partial y^\alpha \quad (\alpha = 1, \dots, n)$$

and the function w is constant. Thus $\partial w / \partial y^\alpha$ ($\alpha = 1, \dots, n$) are zero in the everywhere dense set $V^n - M$ and therefore on the whole hypersurface V^n by continuity. Hence the function w is constant on the whole hypersurface V^n , and the proof of Theorem 1 is complete.

In both parts of Theorem 3 the integral over the boundary V^{n-1} on the left side of the formula (3.13) also vanishes, since over the boundary V^{n-1} $W = 0$ and $N^* = N$ in the two parts respectively. By the same argument as that in the above proof of Theorem 1, we therefore obtain between the two hypersurfaces V^n, V^{*n} a translation, which in part (i) reduces to an identity. Hence Theorem 3 is proved.

Now suppose that in Theorem 3 the fixed direction R is along the vector I_{n+1} and the hypersurfaces V^n, V^{*n} can be represented by equations of the form $y^{n+1} = y^{n+1}(y^1, \dots, y^n)$. Then part (i) of Theorem 3 can be stated as follows: The problem of finding a function $y^{n+1}(y^1, \dots, y^n)$ over a bounded region in the space (y^1, \dots, y^n) with given boundary values such that the first mean curvature M_1 of the hypersurface V^n defined by the equation $y^{n+1} = y^{n+1}(y^1, \dots, y^n)$ is a given function $M_1(y^1, \dots, y^n)$ admits at most one solution. Making use of equations (2.10), (2.13), (4.4) and

$$\frac{\partial g}{\partial x^\alpha} = g g^{\alpha\sigma} \frac{\partial g_{\sigma\sigma}}{\partial x^\alpha},$$

we can easily obtain the first mean curvature of the hypersurface V^n , namely,

$$(4.5) \quad M_1 = n^{-1} g^{-1/2} g^{\alpha\beta} \frac{\partial^2 y^{n+1}}{\partial y^\alpha \partial y^\beta}.$$

Thus the above special case of part (i) of Theorem 3 is a consequence of the well-known uniqueness theorem for the solutions of elliptic differential equations of the second order, since the determinant $|g^{ab}| = 1/g > 0$.

5. Connection with symmetrizations. Let y^1, \dots, y^{n+1} be the coordinates of a point with respect to a fixed orthogonal frame $OI_1 \dots I_{n+1}$ in a Euclidean space E^{n+1} of dimension $n+1 \geq 3$, and let a closed orientable hypersurface V^n twice differentially imbedded in the space E^{n+1} be convex in the direction of the vector I_{n+1} . Let P be any point of the hypersurface V^n , and P^* the other point of intersection of the hypersurface V^n by the line l through the point P and in the direction of the vector I_{n+1} . If the line l is tangent to the hypersurface V^n , then the point P^* coincides with the point P . Let y^{n+1}, y^{*n+1} be respectively the $(n+1)$ th coordinates of the points P, P^* with respect to the frame $OI_1 \dots I_{n+1}$, and M_1^*, N^* the first mean curvature and the unit normal vector of the hypersurface V^n at the point P^* .

The Steiner's symmetrization of the hypersurface V^n with respect to the hyperplane $y^{n+1} = 0$ is a geometric operation by which any point P of the hypersurface V^n goes into a point P' on the line l with

$$y'^{n+1} = \frac{1}{2}(y^{n+1} - y^{*n+1}) = y^{n+1} - \frac{1}{2}(y^{n+1} + y^{*n+1})$$

as its $(n+1)$ th coordinate with respect to the frame $OI_1 \dots I_{n+1}$. In the time interval $0 \leq t \leq 1$, we shift the segment PP^* along its line l into the position $P'P^{*'}$ such that the $(n+1)$ th coordinates of the points $P', P^{*'}$ with respect to the frame $OI_1 \dots I_{n+1}$ are respectively given by

$$(5.1) \quad T_t: y'^{n+1} = y^{n+1} - \frac{t}{2}(y^{n+1} + y^{*n+1}), y^{*n+1} = y^{*n+1} - \frac{t}{2}(y^{n+1} + y^{*n+1}).$$

That is, the segment PP^* is shifted with uniform velocity into the position where it is bisected by the hyperplane $y^{n+1} = 0$. This transformation T_t is called the continuous symmetrization of Steiner.³ T_0 is the identity and T_1 results in a complete symmetrization. It is obvious that the transformation T_t leaves the volume of the hypersurface V^n unchanged.

Now let us consider a neighboring hypersurface $V_{(\epsilon)}^n$ of the hypersurface V^n defined by the vector equation

$$(5.2) \quad Y^{(\epsilon)} = Y + \epsilon(W \cdot N)N,$$

where ϵ is an infinitesimal, Y is the position vector of the point P of the hypersurface V^n with respect to the frame $OI_1 \dots I_{n+1}$, and

$$(5.3) \quad W = wI_{n+1}, \quad w = -y^{n+1} - y^{*n+1}.$$

An elementary calculation and the use of equations (5.2), (2.8), (2.9) yield the coefficients of the first fundamental form of the hypersurface $V_{(\epsilon)}^n$:

³For the continuous symmetrization of Steiner in a Euclidean space E^n of dimension $n = 2, 3$ see (1, pp. 249-251; 4, pp. 200-202).

$$(5.4) \quad g_{\alpha\beta}^{(\epsilon)} = g_{\alpha\beta} - 2\epsilon(W \cdot N) b_{\alpha\beta} + (O)(\epsilon^2),$$

and therefore

$$(5.5) \quad g^{(\epsilon)} = |g_{\alpha\beta}^{(\epsilon)}| = g - 2n \epsilon(W \cdot N) M_1 g + \dots,$$

where the omitted terms are of degrees ≥ 2 in ϵ . From equations (5.5), (2.14) follows immediately the area of the hypersurface $V_{\epsilon}^{(a)}$:

$$(5.6) \quad A^{(\epsilon)} = \int_{V^n} \sqrt{g^{(\epsilon)}} dx^1 \wedge \dots \wedge dx^n = A - n\epsilon \int_{V^n} M_1(W \cdot N) dA + \dots$$

Thus we obtain

$$(5.7) \quad \left(\frac{\partial A^{(\epsilon)}}{\partial \epsilon} \right)_{\epsilon=0} = -n \int_{V^n} M_1(W \cdot N) dA.$$

Similarly, replacing equation (5.2) by $Y^{(\epsilon)} = Y^* + \epsilon(W^* \cdot N^*) N^*$ gives

$$(5.8) \quad \left(\frac{\partial A^{(\epsilon)}}{\partial \epsilon} \right)_{\epsilon=0} = -n \int_{V^n} M_1^*(W^* \cdot N^*) dA^*.$$

Noting that $y^{*n+1} = -y^{n+1} - w$, $W^* = W$ and making use of equation (3.7), we obtain immediately

$$(5.9) \quad W^* \cdot N^* dA^* = -W \cdot N dA,$$

and therefore equation (5.8) becomes

$$(5.10) \quad \left(\frac{\partial A^{(\epsilon)}}{\partial \epsilon} \right)_{\epsilon=0} = n \int_{V^n} M_1^*(W \cdot N) dA.$$

Thus the addition of equations (5.7), (5.10) gives

$$(5.11) \quad \left(\frac{\partial A^{(\epsilon)}}{\partial \epsilon} \right)_{\epsilon=0} = \frac{n}{2} \int_{V^n} (M_1^* - M_1) W \cdot N dA.$$

As in the proof of Theorem 2 in §1, we consider the reflection r with respect to the hyperplane $y^{n+1} = 0$. By this reflection r the point P^* of the hypersurface V^n goes into the point \bar{P}^* defined by

$$(5.12) \quad \bar{Y}^* = Y^* + W,$$

which generates a hypersurface \bar{V}^{*n} . If equation (5.12) is used instead of equation (3.1), then the formula (3.14) becomes

$$(5.13) \quad n \int_{V^n} (M_1^* - M_1) W \cdot N dA + \int_{V^n} (1 - N \cdot \bar{N}^*) (dA + d\bar{A}^*) = 0,$$

where \bar{N}^* and $d\bar{A}^*$ are respectively the unit normal vector and the area element of the hypersurface \bar{V}^{*n} at the point \bar{P}^* . By interchanging the corresponding quantities of the two hypersurfaces V^n , \bar{V}^{*n} at the two points P^* , \bar{P}^* respectively it is easily seen that

$$(5.14) \quad \int_{V^n} (1 - N \cdot \bar{N}^*) d\bar{A}^* = \int_{V^n} (1 - \bar{N}^* \cdot N) dA.$$

By means of equation (5.14), equation (5.13) reduces to

$$(5.15) \quad \frac{n}{2} \int_{V^n} (M_1^* - M_1) W \cdot N \, dA = - \int_{V^n} (1 - N \cdot \bar{N}^*) \, dA,$$

from which and equation (5.11) we therefore obtain

$$(5.16) \quad \left(\frac{\partial A^{(4)}}{\partial \epsilon} \right)_{\epsilon=0} = - \int_{V^n} (1 - N \cdot \bar{N}^*) \, dA.$$

Making use of equations (5.11), (5.15), (5.16) we can easily reach the following conclusion:

If $M_1^* = M_1$ at every point P of the hypersurface V^n , then $(\partial A^{(4)} / \partial \epsilon)_{\epsilon=0} = 0$ and the hypersurface V^n is symmetric with respect to a hyperplane. If the hypersurface V^n is not symmetric with respect to a hyperplane and $\bar{N}^* \neq N$ at every point P of the hypersurface V^n , then $(\partial A^{(4)} / \partial \epsilon)_{\epsilon=0} < 0$.

REFERENCES

1. W. Blaschke, *Vorlesungen über Differentialgeometrie* (Vol. 1, 3rd ed., Berlin, 1930).
2. H. Hopf and K. Voss, *Ein Satz aus der Flächentheorie im Grossen*, Archiv der Mathematik, 3 (1952), 187-192.
3. C. C. Hsiung, *Some integral formulas for closed hypersurfaces*, Math. Scandinavica, 2 (1954), 286-294.
4. G. Pólya and G. Szegő, *Isoperimetric inequalities in mathematical physics* (Ann. Math. Studies, No. 42, Princeton, 1951).

Lehigh University

A NOTE ON THE MATHIEU GROUPS

LOWELL J. PAIGE

1. Introduction. The principal result of this paper is the representation of the Mathieu group M_{23} as a group of 11×11 matrices over the Galois Field $GF(2)$. This is a new representation of M_{23} and in §5 an indication of how the techniques of this result might be extended to the Mathieu group M_{11} is given.

The results of §3 were essentially obtained by Professor E. Spanier while investigating another problem and it was during conversations with him that the present result was observed.

2. Steiner systems and the Mathieu groups. A Steiner system $S(p, q, r)$, with $p < q < r$, is defined on the r integers $1, 2, \dots, r$ and consists of $\binom{r}{q} / \binom{p}{q}$ subsets H_x of q integers each with the property that any arbitrary set of p integers is contained in one and only one of the subsets H_x . For example, the Steiner system $S(2, q+1, q^2+q+1)$ (q a prime) can be constructed by considering the points and lines of a finite projective plane with $q+1$ points on each line.

The group G of a Steiner system $S(p, q, r)$ consists of all those permutations of the symmetric group \mathfrak{S}_r that permute the subsets H_x among themselves. Witt (1, p. 274) has shown that the Steiner systems $S(4, 5, 11)$, $S(5, 6, 12)$, $S(3, 6, 22)$, $S(4, 7, 23)$ and $S(5, 8, 24)$ are unique (i.e., for fixed p, q, r , there exists a permutation of \mathfrak{S}_r carrying $S_1(p, q, r)$ into $S_2(p, q, r)$) and the groups associated with these Steiner systems are the Mathieu groups M_{11} , M_{12} , M_{23} , M_{23} , and M_{24} respectively.

3. Generation of $S(4, 7, 23)$. Let $V(n)$ be the vector space consisting of all n -tuples (x_1, x_2, \dots, x_n) , with each x_i contained in the Galois Field $GF(2)$, under the usual definitions of addition and scalar multiplication.

The distance $d(x, y)$ between two vectors $x = (x_1, x_2, \dots, x_n)$ and $y = (y_1, y_2, \dots, y_n)$ of $V(n)$ is defined to be the number of coordinates for which $x_i \neq y_i$ ($i = 1, 2, \dots, n$).

A subset $S(r, n)$ of $V(n)$ is defined to be an exact r -covering of $V(n)$ if and only if

- (i) For every vector $x \in V(n)$, $\min_{s \in S(r, n)} \{d(x, s)\} \leq r$;
- (ii) For $s_1, s_2 \in S(r, n)$, $d(s_1, s_2) > 2r$.

For a fixed vector s of an exact r -covering $S(r, n)$, the number $N(r, n)$ of vectors $x \in V(n)$ and satisfying $d(x, s) \leq r$ is obviously

Received April 7, 1956.

$$(3.1) \quad N(r, n) = \binom{n}{0} + \binom{n}{1} + \dots + \binom{n}{r}.$$

The number of vectors in an exact r -covering $S(r, n)$ is clearly $2^n/N(r, n)$ since the equation $d(x, s) \leq r$ can have but one solution s in $S(r, n)$ for every vector x of $V(n)$.

It may be possible for $S(r, n)$ to be a linear subspace of $V(n)$. Certainly a necessary condition for such a possibility is that $N(r, n)$ divide 2^n . In the case that $n = 23$ and $r = 3$ we have $N(3, 23) = 2^{11}$ and in the following lemma an exact 3-covering of $V(23)$ is obtained that is a linear subspace.

LEMMA 3.2. *Let R be the subspace of $V(23)$ generated by the rows of the following rectangular array with elements in $GF(2)$:*

$$(3.3) \quad \begin{array}{ccccccccccccccccccccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{array}$$

The linear subspace T orthogonal to R is an exact 3-covering of $V(23)$.

Proof. The crucial part of the proof is the verification that no six columns of (3.3) are linearly dependent. The usual statement at this point regarding straightforward computations would be inappropriate. However, the verification was accomplished on the high speed computer SWAC, and the computations, of necessity, will be omitted.

Assuming the result of the previous paragraph, the proof proceeds by noting that, for any two vectors t_1 and t_2 of T , $d(t_1, t_2) > 6$, since otherwise there would exist a subset of six columns of (3.3) that would be linearly dependent. Now a simple numerical calculation (i.e. $2^{11} \cdot 2^{12} = 2^{23}$) shows that T is an exact 3-covering of $V(23)$.

We now proceed to generate $S(4, 7, 23)$. Let H_t be the set of all integers j such that $t_j \neq 0$ for the vector $t \equiv (t_1, t_2, \dots, t_n)$ of the linear subspace T of Lemma 3.2.

THEOREM 3.4. *The set of all sets H_t containing seven integers forms a Steiner system $S(4, 7, 23)$.*

Proof. T is an exact 3-covering of $V(23)$ and since no six columns of (3.3) are linearly dependent it is clear that every vector of $V(23)$ with 4 non-zero

coordinates must be at distance 3 from those vectors of T with 7 non-zero coordinates. Each vector of T with 7 non-zero coordinates is at distance 3 from $\binom{4}{1}$ vectors with 4 non-zero coordinates and hence there are $\binom{23}{4}/\binom{4}{1} = 253$ vectors in T with 7 non-zero coordinates.

An arbitrary set of 4 integers cannot be contained in H_{t_1} and H_{t_2} (containing 7 integers) if $t_1 \neq t_2$ because this would imply that there would be 6 linearly dependent columns of (3.3). This completes the proof of Theorem 3.4 and establishes the existence of a Steiner system $S(4, 7, 23)$.

4. A matrix representation of M_{23} . In this section an 11×11 matrix representation over $\text{GF}(2)$ will be obtained for the Mathieu group M_{23} .

Let T be the exact 3-covering of $V(23)$ obtained in Lemma 3.2, and let $S(4, 7, 23)$ be the Steiner system consisting of the sets H_i constructed in Theorem 3.4. The following vectors of T are linearly independent and generate T :

$$\begin{aligned} t_1 &= (000000111111100000000000) \\ t_2 &= (111111000000100000000000) \\ t_3 &= (111000111000010000000000) \\ t_4 &= (100110110100001000000000) \\ t_5 &= (010101101100000100000000) \\ t_6 &= (001011011100000010000000) \\ t_7 &= (110010001110000001000000) \\ t_8 &= (101001100110000000100000) \\ t_9 &= (011100010110000000010000) \\ t_{10} &= (100101011010000000001000) \\ t_{11} &= (010011110010000000000010) \\ t_{12} &= (00111010101000000000001) \end{aligned}$$

These vectors yield subsets H_{t_i} ($i = 1, 2, \dots, 12$) of $S(4, 7, 23)$ and since M_{23} consists of those permutations of the symmetric group \mathfrak{S}_{23} that transpose the subsets H_x of $S(4, 7, 23)$ among themselves, it is possible to consider M_{23} as the group of all 23×23 permutation matrices Q which leave T invariant. Thus a representation of M_{23} is induced on the space T and, since each Q is orthogonal, on the space R orthogonal to T .

THEOREM 4.1. *The representation ρ of M_{23} induced on R is an isomorphism.*

Proof. The kernel of ρ consists of those permutations Q that leave the array (3.3) invariant. Since no two columns of (3.3) are the same, the kernel of ρ is the identity.

We thus obtain an 11×11 matrix representation over $\text{GF}(2)$ for M_{23} , and it is of interest to note that this representation is irreducible. The referee has suggested the following simple proof: If ρ were not irreducible, there would exist an invariant subspace S of R of dimension k with $0 < k < 11$. Then $\rho(M_{23})$ would be a subgroup of the group of all non-singular transformations

of R leaving S invariant. However, the order of $\rho(M_{23})$ is divisible by 23 and the order of the group of all non-singular transformations leaving S invariant,

$$2^{k(11-k)} \prod_{i=1}^k (2^k - 2^{i-1}) \prod_{j=1}^{11-k} (2^{11-k} - 2^{j-1}),$$

is not. This argument also proves that M_{23} does not have a faithful representation by $k \times k$ matrices over $\text{GF}(2)$ for any $k < 11$.

5. Comments and generalizations. If the field of coefficients in §3 of $V(n)$ is allowed to be the Galois field $\text{GF}(p^k)$, the same definitions of distance and exact r -covering $S(r, n)$ yield the fact that there are

$$N(r, n, p^k) = \binom{n}{0} + (p^k - 1) \cdot \binom{n}{1} + \dots + (p^k - 1)^r \cdot \binom{n}{r}$$

vectors of $V(n)$ that satisfy $d(x, s) \leq r$ for $s \in S(r, n)$.

In the case that $p = 3, k = 1, r = 2$ we find that $N(2, 11, 3) = 3^5$; as in §3, it is possible to construct a 5×11 rectangular array over $\text{GF}(3)$ such that no 4 columns are linearly dependent. The subsequent analysis of the orthogonal space in $V(11)$ over $\text{GF}(3)$ leads to a Steiner system $S(4, 5, 11)$.

It had been hoped that the matrix representations of the Mathieu group obtained in this paper might lend itself to the determination of a simple set of generators of M_{23} ; unfortunately, this aim has not been realized.

It should also be pointed out that the divisibility of $(p^k)^n$ by $N(r, n, p^k)$, although necessary, is not sufficient to ensure the existence of an exact r -covering of $V(n)$. For example, $N(2, 90, 2) = 2^{12}$, yet a simple analysis of vectors having three non-zero coordinates shows that no exact 2-covering can exist. Here again the techniques of the present note become hopelessly involved in combinatorial analysis if one attempts to find new Steiner systems or simple groups.

REFERENCE

1. E. Witt, *Ueber Steinersche Systeme*, Abhand. Math. Sem. Univ. Hamburg, 12 (1938), 265-275.

University of California at Los Angeles

RELATIVE COHOMOLOGY

D. G. HIGMAN

It is our purpose in this paper to present certain aspects of a cohomology theory of a ring R relative to a subring S , basing the theory on the notions of induced and produced pairs of our earlier paper (2), but making the paper self-contained except for references to a few specific results of (2). The cohomology groups introduced occur in dual pairs. Generic cocycles are defined, and the groups are related to the protraction and retraction of R -modules. Our cohomology groups are modules over the center of R , and in the final section we record some facts concerning their annihilators. Attention is given to the case in which R is a self dual S -ring in the sense of (2). Applications of the theory to the study of orders in algebras will be found in (4), where, in particular, results of (3) are generalized.

Since this paper was first submitted, Professor Hochschild has kindly given the author the opportunity of seeing the manuscript of his paper (8) in which the methods of Cartan-Eilenberg (1) are generalized to give a theory of relative homological algebra. The present paper (as well as (2)) has some results in common with the book of Cartan-Eilenberg and overlaps to some extent with the paper of Hochschild; we have indicated some of the relations in footnotes. Our point of view and methods differ rather widely from Hochschild's.

We are indebted to the referee for suggestions simplifying the notation and increasing the generality somewhat, and for the references to (1).

1. Induced and produced pairs. Let R be a ring with identity element. We shall use the terms *right*, *left* and *two-sided R -module* in the customary way, but always assuming that the identity element of R acts as the identity operator. We shall abbreviate "right R -module" to " R -module."

Let S be a ring with identity element, χ a homomorphism of S into R mapping the identity element of S into that of R . Then every R -module is also an $S\chi$ -module and hence an S -module.

If M is an S -module, the product $M \otimes_S R$ becomes an R -module when one defines $(u \otimes r)x = u \otimes rx$ for $u \in M$, and $r, x \in R$. The \otimes notation is that of (1), $M \otimes_S R$ denoting the tensor product over S of the S -module M with the left S -module R . The pair consisting of this R -module and the natural homomorphism $\kappa: M \rightarrow M \otimes_S R$, which is an S -homomorphism, we shall refer to as the *canonical (R, S, χ) -produced pair determined by M* . We shall omit the (R, S, χ) when no confusion can occur.

The term *(R, S, χ) -produced pair determined by M* will be used to refer to

Received April 3, 1956; in revised form September 8, 1956.

those pairs consisting of an R -module $P(M)$ and an S -homomorphism $\kappa_M: M \rightarrow P(M)$ which satisfy the condition

(P) for each R -module N there exists a homomorphism $\alpha \rightarrow \alpha^*$ of $\text{Hom}_S(M, N)$ into $\text{Hom}_R(P(M), N)$ such that for $\alpha \in \text{Hom}_S(M, N)$, α^* is the unique element of $\text{Hom}_R(P(M), N)$ which makes the diagram

$$\begin{array}{ccc} & P(M) & \\ \kappa_M \uparrow & & \searrow \alpha^* \\ M & \xrightarrow{\alpha} & N \end{array}$$

commutative.¹

The Hom notation is that of (1), $\text{Hom}_S(M, N)$, for example, denoting the module of S -homomorphisms of M into N .

Taking $*$ to be the natural homomorphism of $\text{Hom}_S(M, N)$ into $\text{Hom}_R(M \otimes_S R, N)$, we find that the canonical-produced pair satisfies (I). It follows that any produced pair $(P(M), \kappa_M)$ determined by M is isomorphic with the canonical one, i.e., that there exists an R -isomorphism ϕ of $P(M)$ onto $M \otimes_S R$ such that the diagram

$$\begin{array}{ccc} P(M) & \xrightarrow{\phi} & M \otimes_S R \\ \kappa_M \swarrow & & \nearrow \kappa \\ & M & \end{array}$$

is commutative (2).

The canonical (R, S, χ) -induced pair determined by M consists of the module $\text{Hom}_S(R, M)$, made into an R -module by setting $f^x(r) = f(rx)$ for $f \in \text{Hom}_S(R, M)$ and $r, x \in R$, together with the natural homomorphism $\epsilon: \text{Hom}_S(R, M) \rightarrow M$, which is an S -homomorphism. An (R, S, χ) -induced pair determined by

¹The considerations in (2) were conceived of primarily as generalizations to rings and algebras of certain parts of the theory of representations of finite groups. The term *induced module* was used in connection with $M \otimes_S R$ because of its relation to the classical construction for the induced representation in that theory. This turns out to be unfortunate because produced modules are then injective, induced modules projective, and because of the conflict with the terminology of the representation theory of topological groups. We have therefore switched notation and terminology in the present paper, interchanging "produced" with "induced" and "I" with "I" throughout.

M consists of an R -module $I(M)$ and an S -homomorphism $\epsilon_M: I(M) \rightarrow M$, satisfying the dual of condition (P), namely

(I) for each R -module N there exists a homomorphism $\beta \rightarrow \beta^+$ of $\text{Hom}_S(N, M)$ into $\text{Hom}_R(N, I(M))$ such that for $\beta \in \text{Hom}_S(N, M)$, β^+ is the unique element of $\text{Hom}_R(N, I(M))$ making the diagram

$$\begin{array}{ccc} & I(M) & \\ & \downarrow \epsilon_M & \nearrow \beta^+ \\ M & \xleftarrow{\beta} & N \end{array}$$

commutative.

The canonical-produced pair satisfies (I) with the natural homomorphism

$$\text{Hom}_S(N, M) \rightarrow \text{Hom}_R(N, \text{Hom}_S(R, M))$$

as $+$. Every produced pair determined by M is isomorphic with the canonical one (2).

Henceforth, $(P(M), \kappa_M)$ and $(I(M), \epsilon_M)$ will denote respectively an (R, S, χ) -produced pair and induced pair determined by M . A subscript (R, S) or (R, S, χ) will be used to obtain a more explicit notation when desired; thus $I_{(R, S)}(M)$ for $I(M)$, and so on.

In case M is an S -module by virtue of being an R -module, there exist by (P) and (I) unique R -homomorphisms

$$t_M: P(M) \rightarrow M, j_M: M \rightarrow I(M), \kappa_M t_M = j_M \epsilon_M = 1.$$

Then κ_M and j_M are $1-1$ while t_M and ϵ_M are onto. For the canonical pairs, t_M is the natural homomorphism $M \otimes_S R \rightarrow M$, j_M the natural homomorphism $M \rightarrow \text{Hom}_S(R, M)$. We shall denote by $K(M)$ the kernel of t_M , $K(M) = P(M)(1 - t_M \kappa_M)$, and by $L(M)$ the cokernel of j_M , $L(M) = I(M)/I(M)\epsilon_M j_M$. Thus $K(M)$ and $L(M)$ are R -modules determined up to R -isomorphisms independently of the particular choice of induced and produced pairs. It will be convenient to introduce the notation η_M for the injection $K(M) \rightarrow I(M)$ and π_M for the projection $P(M) \rightarrow L(M)$. Note that there exists a unique S -homomorphism

$$\lambda_M: L(M) \rightarrow P(M), \lambda_M \pi_M = 1, \pi_M \lambda_M = 1 - \epsilon_M j_M.$$

2. The Z_R -modules $\hat{H}^i(M, N)$ and $\hat{H}^i(M, N)$. Given a ring X , Z_X will denote the center of X . If M is an S -module and N an R -module, the module $\text{Hom}_S(M, N)$ becomes a Z_R -module when we define $f^z(u) = f(u)z$ for $f \in \text{Hom}_S(M, N)$, $u \in M$ and $z \in Z$. If M is an R -module, $\text{Hom}_R(M, N)$ is a Z_R -submodule of $\text{Hom}_S(M, N)$. We can verify that the homomorphisms

$$*: \text{Hom}_S(M, N) \rightarrow \text{Hom}_R(P(M), N), \quad +: \text{Hom}_S(N, M) \rightarrow \text{Hom}_R(N, I(M))$$

of conditions (P) and (I) are Z_R -homomorphisms.

Suppose now that M and N are R -modules. We obtain a Z_R -homomorphism $\delta_{M,N}$ of $\text{Hom}_S(M, N)$ into $\text{Hom}_R(K(M), N)$ by following $*$ with the homomorphism of $\text{Hom}_R(P(M), N)$ into $\text{Hom}_R(K(M), N)$ induced by the injection η_M of $K(M)$ into $P(M)$. Thus, for $f \in \text{Hom}_S(M, N)$,

$$f^{\delta_{M,N}} = \eta_M f^*.$$

The kernel of $\delta_{M,N}$ is $\text{Hom}_R(M, N)$. In fact, if f is in this kernel,

$$0 = (1 - t_M \kappa_M) \eta_M f^* = f^* - t_M f.$$

Hence $t_M f = f^*$ is an R -homomorphism, and, since t_M is an R -homomorphism onto, f is an R -homomorphism. On the other hand, if $f \in \text{Hom}_R(M, N)$, $f^* = t_M f$ and

$$f^{\delta_{M,N}} = \eta_M t_M f = 0.$$

Dually, we may define a Z_R -homomorphism $\tilde{\delta}_{M,N}$ of $\text{Hom}_S(N, M)$ into $\text{Hom}_R(N, L(M))$, namely, the product of $+$ with the homomorphism of $\text{Hom}_R(N, I(M))$ into $\text{Hom}_R(N, L(M))$ induced by π_M :

$$g^{\tilde{\delta}_{M,N}} = g^+ \pi_M$$

for $g \in \text{Hom}_S(N, M)$. The kernel of $\tilde{\delta}_{M,N}$ is $\text{Hom}_R(N, M)$.

The R -module M determines R -modules $K^i(M)$ and $P^i(M)$ ($i = 0, 1, \dots$), defined by the recursive formulas

$$K^0(M) = P^0(M) = M, \quad K^{i+1}(M) = K(K^i(M)), \quad P^{i+1}(M) = P(K^i(M)).$$

Dually, M determines R -modules $L^i(M)$ and $I^i(M)$, ($i = 0, 1, \dots$), according to²

$$L^0(M) = I^0(M) = M, \quad L^{i+1}(M) = L(L^i(M)), \quad I^{i+1}(M) = I(L^i(M)).$$

We shall now define a Z_R -complex $(C(M, N), \delta)$, determined by the ordered pair M, N of R -modules, by letting

$$C^i(M, N) = \text{Hom}_S(K^i(M), N), \quad \delta^i = \delta_{K^i(M), N} \quad \text{for } i \geq 0,$$

and

$$C(M, N) = (0), \quad \delta = 0 \text{ for } i < 0.$$

We have $\delta^{i-1} \delta^i = 0$ for all i , since, for $i \geq 0$, the image $B^i(M, N)$ of δ^{i-1} is contained in $\text{Hom}_R(K^i(M), N)$, which is the kernel of δ^i . The cohomology groups of this complex, which are Z_R -modules, and which are determined up to Z_R -isomorphism independently of the particular choice of induced pair determined by M , we shall denote by $H^i(M, N)$. More explicit notation

²This amounts to constructing the standard (R, S) -projective and injective resolutions of M as in (8).

such as $H^i_{(R,S)}(M, N)$ will be used where desirable.² It is immediate that $H^0(M, N) = \text{Hom}_R(M, N)$, and, for $i > 0$, $\alpha \geq 0$,

$$H^{i+\alpha}(M, N) = H^i(K^\alpha(M), N).$$

We may dualize the above construction to obtain a second Z_R -complex $(\tilde{C}(M, N), \tilde{\delta})$, taking

$$\tilde{C}^i(M, N) = \text{Hom}_S(N, L^i(M)), \quad \tilde{\delta}^i = \tilde{\delta}_{L^i(M), N}, \quad i \geq 0,$$

and $\tilde{C}^i(M, N) = (0)$, $\tilde{\delta}^i = 0$ for $i < 0$. We shall denote the image of δ^{i-1} by $\tilde{B}^i(M, N)$, and the cohomology groups of this complex by $\tilde{H}^i(M, N)$. These are again Z_R -modules, determined up to Z_R -isomorphism independently of the choice of produced pair determined by M . We have

$$\tilde{H}^0(M, N) = \text{Hom}_R(N, M), \quad \tilde{H}^{i+\alpha}(M, N) = \tilde{H}^i(L^\alpha(M), N)$$

for all $i > 0$, $\alpha \geq 0$.

3. The isomorphism $\tilde{\sigma}$. Let M and N be S -modules. There is a $Z_R \cap S_X$ -isomorphism σ of $\text{Hom}_S(P(M), N)$ onto $\text{Hom}_S(M, I(N))$ mapping

$$f \in \text{Hom}_S(P(M), N)$$

onto $f^\sigma = \kappa_M f^+$. The inverse τ of σ maps $g \in \text{Hom}_S(M, I(N))$ onto $g^\tau = g^* \epsilon_N$. In fact, $(\kappa_M f^+)^* = f^+$, and hence $f^{\sigma\tau} = (\kappa_M f^+)^* \epsilon_N = f^+ \epsilon_N = f$, so that $\sigma\tau = 1$. Dually, $\tau\sigma = 1$. In the case of the canonical-induced and -produced pairs, σ is the natural homomorphism of $\text{Hom}_S(M \otimes_S R, N)$ onto $\text{Hom}_S(M, \text{Hom}_S(R, N))$.

Now assume that M and N are R -modules so that $K(M)$ and $L(N)$ are defined. We then have a $Z_R \cap S_X$ -isomorphism $\tilde{\sigma}$ of $\text{Hom}_S(K(M), N)$ onto $\text{Hom}_S(M, L(N))$, mapping $f \in \text{Hom}_S(K(M), N)$ onto

$$f^{\tilde{\sigma}} = [(1 - t_M \kappa_M) f]^{\sigma} \pi_N.$$

Inverse to $\tilde{\sigma}$ is $\tilde{\tau}$ defined by

$$g^{\tilde{\tau}} = \eta_M (g \lambda_N)^{\tau}$$

for $g \in \text{Hom}_S(M, L(N))$. Indeed,

$$\begin{aligned} f^{\tilde{\sigma}} \lambda_N &= \kappa_N [(1 - t_M \kappa_M) f]^+ \pi_N \lambda_N = \kappa_M [(1 - t_M \kappa_M) f]^+ (1 - \epsilon_N j_N) \\ &= \kappa_M [(1 - t_M \kappa_M) f]^+ - \kappa_M (1 - t_M \kappa_M) f j_N = [(1 - t_M \kappa_M) f]^{\sigma}. \end{aligned}$$

Hence

$$f^{\tilde{\sigma}\tilde{\tau}} = \eta_M (f^{\tilde{\sigma}} \lambda_N)^{\tau} = \eta_M [(1 - t_M \kappa_M) f]^{\sigma\tau} = f,$$

whence $\tilde{\sigma}\tilde{\tau} = 1$. Dually, $\tilde{\tau}\tilde{\sigma} = 1$.

²It can easily be proved that $H^i_{(R,S)}(M, N)$ is isomorphic with $\text{Ext}^i_{(R,S_X)}(M, N)$ as defined by Hochschild (8). When this is taken into consideration, the connection between the results of §§2-7 of the present paper and (8) will be seen.

We shall show that the diagram

$$\begin{array}{ccc} \text{Hom}_S(M, N) & \xrightarrow{\delta_{MN}} & \text{Hom}_S(K(M), N) \\ & \searrow \delta_{N, M} & \downarrow \tilde{\sigma} \\ & & \text{Hom}_S(M, L(N)) \end{array}$$

is anti-commutative. If $f \in \text{Hom}_S(M, N)$,

$$(1 - \iota_M \kappa_M) f^{\tilde{\delta}} = (1 - \iota_M \kappa_M) \eta_M f^* = f^* - \iota_M f.$$

Hence

$$f^{\tilde{\delta} \tilde{\sigma}} = (f^* - \iota_M f)^{\sigma} \pi_N = f^{\sigma} \pi_N - (\iota_M f)^{\sigma} \pi_N.$$

But

$$f^{\sigma} \pi_N = \kappa_M (f^*)^+ \pi_N = \kappa_M f^* j_N \pi_N = 0,$$

and

$$(\iota_M f)^{\sigma} \pi_N = \kappa_M (\iota_M f)^+ \pi_N = \kappa_M \iota_M f^+ \pi_N = f^+ \pi_N = f^{\tilde{\delta}}.$$

Hence

$$f^{\tilde{\delta} \tilde{\sigma}} = -f^{\tilde{\delta}}.$$

Moreover, the diagram

$$\begin{array}{ccc} \text{Hom}_S(K(M), N) & & \\ \tilde{\sigma} \downarrow & \searrow \delta_{N, K(M)} & \\ \text{Hom}_S(M, L(N)) & \xrightarrow{\delta_{M, L(N)}} & \text{Hom}_S(K(M), L(N)) \end{array}$$

is commutative. For, if $f \in \text{Hom}_S(K(M), N)$,

$$f^{\tilde{\sigma}} = \kappa_{K(M)} [(1 - \iota_{K(M)} \kappa_{K(M)}) f]^+ \pi_N,$$

hence

$$(f^{\tilde{\sigma}})^* = [(1 - \iota_{K(M)} \kappa_{K(M)}) f]^+ \pi_N.$$

Thus

$$\begin{aligned} f^{\tilde{\sigma} \tilde{\delta}} &= \eta_{K(M)} (f^{\tilde{\sigma}})^* = \eta_{K(M)} [(1 - \iota_{K(M)} \kappa_{K(M)}) f]^+ \pi_N \\ &= \eta_{K(M)} [(1 - \iota_{K(M)} \kappa_{K(M)}) f]^+ \pi_N = f^+ \pi_N = f^{\tilde{\delta}}, \end{aligned}$$

proving the desired result.

Applying these results to the complexes $(C(M, N), \delta)$ and $(\bar{C}(N, M), \bar{\delta})$ we obtain an anti-commutative diagram

$$\begin{array}{ccccccc} \dots & \rightarrow & C^{-1}(M, N) & \xrightarrow{\delta^{-1}} & C^0(M, N) & \xrightarrow{\delta^0} & C^1(M, N) \xrightarrow{\delta^1} C^2(M, N) \rightarrow \dots \\ & & \parallel & & \parallel & & \downarrow \bar{\sigma} \qquad \downarrow \bar{\sigma} \\ \dots & \rightarrow & \bar{C}^{-1}(N, M) & \xrightarrow{\bar{\delta}^{-1}} & \bar{C}^0(N, M) & \xrightarrow{\bar{\delta}^0} & \bar{C}^1(N, M) \xrightarrow{\bar{\delta}^1} \bar{C}^2(N, M) \rightarrow \dots \end{array}$$

Consequently

THEOREM 1. $\bar{\sigma}$ induces a $Z_R \cap S_X$ isomorphism of $H^i(M, N)$ onto $\bar{H}^i(N, M)$ for $i \geq 0$.

COROLLARY. $H^{i+\alpha}(M, N) \simeq H^i(M, L^\alpha(N))$ and $\bar{H}^{i+\alpha}(M, N) \simeq \bar{H}^i(M, K^\alpha(N))$ for $i > 0, \alpha \geq 0$, the isomorphisms being $Z_R \cap S_X$ -isomorphisms.

4. Some special induced and produced pairs. To give explicit constructions for the cohomology groups introduced above one need only supply particular induced and produced pairs, the canonical ones not always being the most suitable. Thus, for example, suppose that B and A are rings, and let T be a subring of A containing the identity element thereof. Let M be a $B' \otimes T$ -module. Here the $'$ indicates mirror image, and \otimes the tensor product over the ring of the rational integers. A $B' \otimes T$ -module may be considered as a left B -, and a right T -module, and we shall use corresponding notation where convenient. The module $M \otimes_T A$ becomes a $B' \otimes A$ -module when we let $b(u \otimes a) = bu \otimes a$ and $(u \otimes a)x = ax$ for $b \in B, u \in M$ and $a, x \in A$, while $\text{Hom}_T(A, M)$ becomes a $B' \otimes A$ -module when we let ${}^b f(a) = b[f(a)]$ and $f^a(x) = f(xa)$ for $b \in B, f \in \text{Hom}_T(A, M)$ and $x, a \in A$. As may be seen by verifying (I) and (P), combining these modules with the natural homomorphisms $M \rightarrow M \otimes_T A$ and $\text{Hom}_T(A, M) \rightarrow M$ yields respectively a $(B' \otimes A, B' \otimes T, \chi)$ -induced and produced pair determined by M , where χ is the natural homomorphism $B' \otimes T \rightarrow B' \otimes A$. If M is a $B' \otimes A$ -module, $K(M)$ and $L(M)$ are the respective kernels of the natural homomorphisms $M \otimes_T A \rightarrow M$ and $M \rightarrow \text{Hom}_T(A, M)$.

Now let M be a $B' \otimes T$ -module, N a $B' \otimes A$ -module. Further, let U be a subring of $T \cap Z_A$, containing the identity element of A . The module $\text{Hom}_{B' \otimes U}(M, N)$ attains the status of a $T' \otimes_U A$ -module when we define

$${}^t f(u) = f(ut) \text{ and } f^x(u) = f(ux)$$

for

$$f \in \text{Hom}_{B' \otimes U}(M, N), \quad t \in T, x \in A, u \in M;$$

let us denote it by $\Phi(M, N)$. If M is a $B' \otimes A$ -module we may replace t above by $y \in A$, turning $\Phi(M, N)$ into an $A' \otimes_U A$ -module. We shall outline how one may prove⁴

⁴Combining Theorems 1 and 2 gives Theorem 2 of (8).

THEOREM 2. If M and N are $B' \otimes A$ -modules, then there exists a natural Z_A -isomorphism of

$$H^i(B' \otimes A, B' \otimes T)(M, N) \text{ onto } H^i(C, D)(\Phi(M, N), A),$$

where $C = A' \otimes_U A$, $D = T' \otimes_U A$.

The first of the cohomology groups mentioned is a $Z_{B' \otimes A}$ -module, the second a Z_C -module, so it makes sense to speak of a Z_A -isomorphism between them. The natural homomorphism $\eta: D \rightarrow C$ is understood.

The main step in the proof consists in the construction of a suitable (C, D, η) -induced pair for $\Phi(M, N)$. To this end, consider the diagram

$$\begin{array}{ccc} \Phi(P(M), N) & & \\ \epsilon \downarrow & \nwarrow \beta^+ & \\ \Phi(M, N) & \xleftarrow{\beta} & H \end{array}$$

where (a) ϵ is the D -homomorphism induced by $\kappa_M: M \rightarrow P(M)$, $(P(M), \kappa_M)$ being a $(B' \otimes A, B' \otimes T, \chi)$ -produced pair determined by M , and (b) β is a D -homomorphism of the C -module H . If now β^+ is a C -homomorphism such that the diagram is commutative, then for $h \in H$, $u \in M$, $(h\beta^+)(u\kappa_M) = (h\beta)(u)$, hence for $a \in A$,

$$(h\beta^+)(u\kappa_M a) = (h\beta^+)(u\kappa_M) = [(ah)\beta](u),$$

i.e.,

$$4.1 \quad (h\beta^+)(u\kappa_M a) = [(ah)\beta](u).$$

On the other hand, we may verify that the formula 4.1 does indeed define a C -homomorphism making the diagram commutative. Consequently the pair $(\Phi(P(M), N), \epsilon)$ is a (C, D, η) -induced pair for $\Phi(M, N)$.

If M is a $B' \otimes A$ -module, $\Phi(M, N)$ is a C -module, and the homomorphism $\iota_M: P(M) \rightarrow M$ induces the unique $A' \otimes_U A$ -homomorphism $j: \Phi(M, N) \rightarrow \Phi(P(M), N)$ such that $j\epsilon = 1$, as is seen by taking $\beta = 1$ and $H = \Phi(M, N)$ in 4.1. Consequently the sequence

$$(0) \rightarrow \Phi(M, N) \xrightarrow{j} \Phi(P(M), N) \xrightarrow{\gamma} \Phi(K(M), N) \rightarrow (0)$$

is exact, where γ is induced by the injection $K(M) \rightarrow P(M)$. In the complex for constructing the groups $H^i(\Phi(M, N), A)$ from the produced pair $(\Phi(P(M), N), \epsilon)$ we may therefore identify $L(\Phi(M, N))$ with $\Phi(K(M), N)$ and the projection of $\Phi(P(M), N)$ onto $L(\Phi(M, N))$ with γ .

The modules $\text{Hom}_{B' \otimes T}(M, N)$ and $\text{Hom}_D(A, \Phi(M, N))$ are in particular Z_A -modules, the first being a $Z_{B' \otimes A}$ - and the second a Z_C -module. The

natural isomorphism between them is a Z_A -isomorphism. One may now verify the commutativity of the diagram

$$\begin{array}{ccc} \text{Hom}_{B' \otimes T}(M, N) & \rightarrow & \text{Hom}_{B' \otimes T}(K(M), N) \\ \downarrow & & \downarrow \\ \text{Hom}_D(A, \Phi(M, N)) & \rightarrow & \text{Hom}_D(A, \Phi(K(M), N)) \end{array}$$

where (1) the arrows pointing down represent the natural isomorphisms, (2) the top arrow represents $\delta_{M, N}$, and (3) the bottom arrow represents the product of

$$+ : \text{Hom}_D(A, \Phi(M, N)) \rightarrow \text{Hom}_D(A, \Phi(P(M), N))$$

with the homomorphism of the second of these modules into

$$\text{Hom}_D(A, \Phi(K(M), N))$$

induced by γ . Application of this fact to the appropriate complexes gives Theorem 2.

5. Generic cocycles. Let M, N and X be R -modules, and let

$$f \in \text{Hom}_S(K(M), N).$$

Then $\mu_f: g \rightarrow gf$ for $g \in \text{Hom}_S(X, K(M))$ defines a $Z_R \cap S_X$ -homomorphism μ_f of $\text{Hom}_S(X, K(M))$ into $\text{Hom}_S(X, N)$. If

$$f \in \text{Hom}_R(K(M), N)$$

we see that μ_f is a Z_R -homomorphism, and moreover, that the diagram

$$\begin{array}{ccc} \text{Hom}_S(X, K(M)) & \xrightarrow{\delta_{X, K(M)}} & \text{Hom}_S(K(X), K(M)) \\ \downarrow \mu_f & & \downarrow \mu_f \\ \text{Hom}_S(X, N) & \xrightarrow{\delta_{X, N}} & \text{Hom}_S(K(X), N) \end{array}$$

is commutative. For then

$$g^{\mu_f \delta} = \eta_X(gf)^* = \eta_X g^* f = g^{\delta \mu_f}.$$

Application of this to the complexes $(C(X, K(M)), \delta)$ and $(C(X, N), \delta)$ proves that μ_f induces a Z_R -homomorphism of $H^\alpha(X, K(M))$ into $H^\alpha(X, N)$ for all $\alpha \geq 0$. In particular:

If $f \in \text{Hom}_S(K^i(M), N)$, μ_f induces a Z_R -homomorphism of $H^\alpha(M, K^i(M))$ into $H^\alpha(M, N)$ for all $\alpha \geq 0$, and $\mu_f: I^i \rightarrow f$, where I^i is the identity automorphism of $K^i(M)$.

We shall refer to the element $I^i \in B^i(M, K^i(M))$ as the first generic i -cocycle determined by M .

Dually, if

$$f \in \text{Hom}_S(N, L(M)), \quad \lambda_f: g \rightarrow fg$$

defines a $Z_R \cap S_X$ -homomorphism λ_f of $\text{Hom}_S(L(M), X)$ into $\text{Hom}_S(N, X)$,

which is a Z_R -homomorphism if $f \in \text{Hom}_R(N, L(M))$. In the latter case, the diagram

$$\begin{array}{ccc} \text{Hom}_S(L(M), X) & \xrightarrow{\tilde{\delta}_{X, L(M)}} & \text{Hom}_S(L(M), L(X)) \\ \lambda_f \downarrow & & \lambda_f \downarrow \\ \text{Hom}_S(N, X) & \xrightarrow{\tilde{\delta}_{X, N}} & \text{Hom}_S(N, L(X)) \end{array}$$

is commutative. Hence λ_f induces a Z_R -homomorphism of $H^\alpha(X, L(M))$ into $H^\alpha(X, N)$, $\alpha \geq 0$. In particular:

If $f \in \text{Hom}(N, L^i(M))$, λ_f induces a Z_R -homomorphism of $H^\alpha(N, L^i(M))$ into $H^\alpha(M, N)$ for all $\alpha \geq 0$, and $\lambda_f: J^i \rightarrow f$, where J^i is the identity automorphism of $L^i(M)$.

We shall refer to the element $J^i \in \hat{B}^i(M, K^i(M))$ as the *second generic i-cocycle determined by M*.

As a consequence of the above considerations we obtain

THEOREM 3. *If i is a positive integer, and M is an R -module, then the following conditions imply each other.*

- (a) $H^i(M, N) = (0)$ for all R -modules N .
- (b) $H^i(M, K^i(M)) = (0)$.
- (c) $J^i \in B^i(M, N)$.
- (d) $H^{i+\alpha}(M, N) = (0)$ for all R -modules N and all $\alpha \geq 0$.

The dual Theorem 3' is obtained by replacing H with \hat{H} , K with L , B with \hat{B} , and I with J .

6. Protractions and retractions.⁵ Let M be an R -module. If H is an R -module and $\phi: M \rightarrow H$ is an R -homomorphism, we shall call the pair (H, ϕ) and (R, S) -protraction of M provided that there exists an S -homomorphism $\lambda: M \rightarrow H$ such that $\lambda\phi = 1$. The kernel $N = H(1 - \phi\lambda)$ of ϕ will be called the *kernel of (H, ϕ)* . Two (R, S) -protractions (H_1, ϕ_1) and (H_2, ϕ_2) of M with kernel N will be called *R -isomorphic* if there exists an R -isomorphism μ of H_1 onto H_2 such that $\phi_1 = \mu\phi_2$.

Corresponding to an (R, S) -protraction of M with kernel N there is an element $f_\lambda \in \text{Hom}_R(K(M), N)$ defined by

$$f_\lambda = \eta_M \lambda^*(1 - \phi\lambda).$$

It can be seen that the correspondence $(H, \phi) \rightarrow f_\lambda$ induces a 1-1 mapping of the set of classes of R -isomorphic (R, S) -protractions of M with kernel N onto $H^i(M, N)$, which becomes a group isomorphism when the Baer composition is introduced into the set of classes of protractions.

The (R, S) -protraction (H, ϕ) of M with kernel N is said to *split* if there exists an R -homomorphism $\alpha: M \rightarrow H$ such that $\alpha\phi = 1$. Two R -isomorphic protractions split, or do not, together. Let $f_\lambda \in \text{Hom}_R(K(M), N)$ correspond

⁵The material of this section overlaps considerably with Cartan and Eilenberg (1, §6, Ch. II), as well as with Hochschild (8).

to (H, ϕ) as above. Then it is easy to see that each of the following conditions is necessary and sufficient for (H, ϕ) to split:

- (a) there exists an R -submodule N^* of H such that $H = N \oplus N^*$,
- (b) $H \simeq N \oplus M$ as an R -module,
- (c) $f_\lambda \in B^1(M, N)$.

The pair $(I(M), t_M)$ is an (R, S) -protraction of M with kernel $K(M)$, and corresponds to the class of the first generic 1-cocycle

$$I^1 = \int_{\kappa_M}$$

An R -module M will be called (R, S) -projective if, whenever (H', ϕ) is an (R, S) -protraction of an R -module H and $\alpha: M \rightarrow H$ is an R -homomorphism, there exists an R -homomorphism $\bar{\alpha}: M \rightarrow H$ such that $\alpha = \bar{\alpha}\phi$. From (2, Theorem 6) Theorem 3, and the above remarks we conclude

THEOREM 4. *Each of the following conditions is necessary and sufficient for an R -module M to be (R, S) -projective.*

- (a) *The (R, S) -protraction $(P(M), t_M)$ splits.*
- (b) *Every (R, S) -protraction of M splits.*
- (c) *$H^1(M, N) = (0)$ for all R -modules N .*

If U is an S -module, the R -module $P(U)$ is (R, S) -projective according to (2, Theorem 3). Hence we have

COROLLARY. *If U is an S -module, $H^i(P(U), N) = (0)$ for all R -modules N and all $i > 0$.*

A pair (H, ψ) consisting of an R -module H and an R -homomorphism $\psi: M \rightarrow H$ will be called an (R, S) -retraction of M with kernel N if there exists an S -homomorphism $\mu: H \rightarrow M$ such that $\psi\mu = 1$, and if N is the cokernel of ψ , $N = H/M\psi$. This is the dual of the concept of (R, S) -protraction. We define R -isomorphism between retractions by dualizing the corresponding concept for protractions, and obtain a 1-1 correspondence between the set of classes of isomorphic (R, S) -retractions of M with cokernel N and the elements of $\bar{H}^1(M, N)$ (and hence of $H^1(N, M)$ by Theorem 1). The definition of splitting for (R, S) -retractions is dual to that for protractions. Of course there is a 1-1 correspondence between the set of (R, S) -protractions of M with kernel N and the (R, S) -retractions of N with cokernel N , such that a protraction splits if and only if the corresponding retraction splits.

$(P(M), j_M)$ is an (R, S) -retraction of M with cokernel $L(M)$, and corresponds to the class of the second generic 1-cocycle J^1 .

Dual to (R, S) -projective modules we have (R, S) -injective modules, and dual to Theorem 4 we have

THEOREM 4'. *The following conditions*

- (a) *The (R, S) -retraction $(I(M), j_M)$ of M splits.*
- (b) *Every (R, S) -retraction of M splits.*
- (c) *$\bar{H}^1(M, N) = (0)$ for all R -modules N .*

are each necessary and sufficient for an R -module M to be (R, S) -injective.

If U is an S -module, $I(U)$ is (R, S) -injective according to (2, Theorem 3'). Consequently

COROLLARY. For every S -module U and every R -module N , $\hat{H}^i(I(U), N) = (0)$ for all $i > 0$.

7. Cohomology dimension. If an R -module M is (R, S) -projective [injective], then according to (2, Theorem 6) so is $K(M)$ [$L(M)$]. Under certain circumstances the converse is true. We shall consider the hypothesis $(R, S; M)$ There exists an R -isomorphism μ_M of $I(M)$ onto $P(M)$. If $(R, S; M)$ holds, then M is (R, S) -projective if and only if it is (R, S) -injective, as follows from (2, Theorems 6, 6').

THEOREM 5. Suppose that the hypothesis $(R, S; K(M))$ holds. Then $K(M)$ (R, S) -projective implies that M is (R, S) -projective.

Proof. If $(R, S; K(M))$ holds and $K(M)$ is (R, S) -projective then $K(M)$ is (R, S) -injective. Hence the (R, S) -retraction $(P(M), \eta_M)$ of $K(M)$ splits by (2, Theorem 6'). Hence $P(M) \simeq M \otimes K(M)$ as an R -module, so that M is (R, S) -projective by (2, Theorem 6).

The dual Theorem 5' is obtained by replacing K with L and projective with injective.

It will be convenient to denote by $d_{(R, S)} M$ the smallest integer $i > 0$ such that $H^i(M, N) = (0)$ for all R -modules N , if such an i exists, setting $d_{(R, S)} M = \infty$ otherwise. The dual $\bar{d}_{(R, S)} M$ is defined by replacing H by \hat{H} . By Theorems 4, 4', $d_{(R, S)} M \leq i$ if and only if $K^{i-1}(M)$ is (R, S) -projective, while $\bar{d}_{(R, S)} M \leq i$ if and only if $L^{i-1}(M)$ is (R, S) -injective. By Theorem 5, if $(R, S; K^i(M))$ holds, then $d_{(R, S)} M \leq i + 1$ implies $d_{(R, S)} M \leq i$, while if $(R, S; L^i(M))$ holds, then $\bar{d}_{(R, S)} M \leq i + 1$ implies $\bar{d}_{(R, S)} M \leq i$.

The two conditions

$$(c.i) \quad d_{(R, S)} M \leq i \text{ for all } R\text{-modules } M.$$

$$(c'.i) \quad \bar{d}_{(R, S)} M \leq i \text{ for all } R\text{-modules } M.$$

imply each other as we deduce at once from Theorem 1. We define class (R, S) to be the minimum integer $i > 0$ such that (c.i) and (c'.i) hold if such an i exists, letting class $(R, S) = \infty$ otherwise. From the above we have

THEOREM 6. If the hypothesis $(R, S; M)$ holds for every R -module M , then class $(R, S) < \infty$ implies class $(R, S) = 1$.

The hypotheses of this theorem are satisfied if R is a self dual S -ring in the sense of (2).

Now let us suppose that B, A, T and U are rings as in §4, and let M and N be $B' \otimes A$ -modules. If

$$d_{(B' \otimes A, B' \otimes T)} M = i < \infty,$$

then $I^i(M) \simeq K^i(M) \otimes K^{i+1}(M)$ as a $B' \otimes A$ -module, according to Theorem 6 of (2). Consequently

$$\Phi(I^i(M), N) \simeq \Phi(K^i(M), N) \oplus \Phi(K^{i+1}(M), N)$$

as an $A' \otimes_v A$ -module, whence one concludes by the construction of §4 that

$$d_{(A' \otimes_v A, T' \otimes_v A)}(M, N) \leq i.$$

In particular, if M is $(B' \otimes A, B' \otimes T)$ -projective then $\Phi(M, N)$ is $(A' \otimes_v A, T' \otimes_v A)$ -injective (8, Lemma 2). Moreover, we conclude that

$$\begin{aligned} \text{class}(A' \otimes_v A, T' \otimes_v A) &= d_{(A' \otimes_v A, T' \otimes_v A)} A = \bar{d}_{(A' \otimes_v A, T' \otimes_v A)} A \\ &> \text{class}(B' \otimes A, B' \otimes T), \end{aligned}$$

considering A in the natural way as an $A' \otimes_v A$ -module.

8. The ideals $\mathfrak{I}^i(M)$ and $\bar{\mathfrak{I}}^i(M)$. The results of §5 can be refined as follows. We shall denote by $\mathfrak{I}^i(M, N)$ [or $\mathfrak{I}^i_{(R, S)}(M, N)$] the annihilator of the Z_R -module $H^i(M, N)$. Letting $\mathfrak{I}^i(M)$ denote the intersection over all R -modules N of the ideals $\mathfrak{I}^i(M, N)$ we have by §5 that

$$\begin{aligned} \mathfrak{I}^i(M) &= \mathfrak{I}^i(M, K^i(M)) \\ &= \{\omega \in Z_R \mid \zeta\omega \in \bar{B}^i(M, K^i(M))\}. \end{aligned}$$

Here $\zeta\omega$ denotes right operation by ω ; $\zeta\omega: u \rightarrow u\omega$ for $u \in K^i(M)$, $\omega \in Z_R$, so that $\zeta\omega = I^\omega$. Condition (a) of Theorem 3 is equivalent to the condition that $\mathfrak{I}^i(M) = Z_R$.

Dually, we define $\bar{\mathfrak{I}}^i(M, N)$ to be the annihilator of the Z_R -module $\bar{H}^i(M, N)$, and $\bar{\mathfrak{I}}^i(M)$ to be the intersection over all R -modules N of these ideals. Then

$$\begin{aligned} \bar{\mathfrak{I}}^i(M) &= \bar{\mathfrak{I}}(M, L^i(M)) \\ &= \{\omega \in Z_R \mid \zeta\omega \in \bar{B}^i(M, L^i(M))\}. \end{aligned}$$

The condition dual to condition (a) of Theorem 3 is equivalent to the condition that $\bar{\mathfrak{I}}^i(M) = Z_R$.

We have at once that for $i > 0$, $\alpha > 0$,

$$\mathfrak{I}^{i+\alpha}(M, N) = \mathfrak{I}^i(K^\alpha(M), N), \quad \bar{\mathfrak{I}}^{i+\alpha}(M, N) = \bar{\mathfrak{I}}^i(L^\alpha(M), N).$$

Moreover, by Theorem 1,

$$\mathfrak{I}^i(M; N) \cap S_\alpha = \bar{\mathfrak{I}}^i(N, M) \cap S_\alpha$$

while the Corollary to Theorem 1 implies that

$$\mathfrak{I}^i(M) \cap S_\alpha \subseteq \mathfrak{I}^{i+\alpha}(M), \quad \bar{\mathfrak{I}}^i(M) \cap S_\alpha \subseteq \bar{\mathfrak{I}}^{i+\alpha}(M)$$

for $i > 0$, $\alpha > 0$.

Let (H, ϕ) be an (R, S) -protraction of the R -module M with kernel N (§6). Then there exists an S -homomorphism $\lambda: M \rightarrow H$ such that $\lambda\phi = 1$, and $N = H(1 - \phi\lambda)$. The corresponding element $f_\lambda \in \text{Hom}_R(M, N)$ is defined by $f_\lambda = \eta_M \lambda^*(1 - \phi\lambda)$.

PROPOSITION. *If $\omega \in Z_R$, then $f_\lambda \omega \in B^1(M, N)$ if and only if there exists an R -homomorphism $\beta: M \rightarrow H$ such that $\beta\phi = \zeta\omega$, where $\zeta\omega$ denotes right operation on M by ω .*

Proof. Suppose that

$$f_\lambda \omega = g^{\delta_{M,N}}, \quad g \in \text{Hom}_S(M, N).$$

This means that $\eta_M \lambda^*(1 - \phi\lambda)\zeta\omega = \eta_M g^*$. Then, if η is the injection $N \rightarrow H$,

$$\begin{aligned} 0 &= \eta_M [\lambda^*(1 - \phi\lambda)\zeta\omega - \eta_M g^*] \eta = \eta_M [\lambda^* \zeta\omega - g^* \eta] \\ &= \eta_M [\lambda \zeta\omega - g \eta]^* = [\lambda \zeta\omega - g \eta]^{\delta_{M,N}}. \end{aligned}$$

Consequently $\beta = \lambda \zeta\omega - g \eta$ is an element of $\text{Hom}_R(M, H)$. Further, $\beta\phi = \lambda \zeta\omega\phi - g \eta\phi = \zeta\omega$.

Suppose on the other hand that there exists an R -homomorphism $\beta: M \rightarrow H$ such that $\beta\phi = \zeta\omega$. Since

$$\eta_M \beta^* = \eta_M \beta t_M = 0,$$

if we let $\gamma = \lambda \zeta\omega - \beta$, we have

$$(f_\lambda \omega) \eta = \eta_M (\lambda^* \zeta\omega - \beta^*) = \eta_M \gamma^* = g^{\delta_{M,N}} \eta,$$

where $g = \gamma(\zeta\omega - \phi\gamma)$ is an element of $\text{Hom}_S(M, N)$. Hence

$$f_\lambda \omega = g^{\delta_{M,N}}.$$

There is the dual result for (R, S) -retractions of M .

Application of this proposition to the (R, S) -protraction $(P(M), t_M)$ of M gives (a) of the following theorem, (a') being its dual.

THEOREM 7. *If $\omega \in Z_R$, then*

(a) $\omega \in \mathfrak{I}^1(M)$ if and only if there exists $\beta \in \text{Hom}_R(M, P(M))$ such that $\beta t_M = \zeta\omega$.

(a') $\omega \in \overline{\mathfrak{I}}^1(M)$ if and only if there exists $\beta \in \text{Hom}_R(I(M), M)$ such that $j_M \beta = \zeta\omega$.

$$\begin{array}{ccc} I(M) & \xrightarrow{j_M} & P(M) \\ j_M \uparrow & \swarrow \alpha^* & \downarrow t_M \\ M & \xleftarrow{\alpha} & M \end{array}$$

Assuming hypothesis $(R, S; M)$ of §7, namely, the existence of an R -isomorphism μ_M of $I(M)$ onto $P(M)$, we may construct the Casimir operators as in (2). Thus, if $\alpha \in \text{Hom}_S(M, M)$, $c(\alpha) = \alpha^+ \mu_M t_M$ and $\bar{c}(\alpha) = j_M \mu_M \alpha^*$ are elements of $\text{Hom}_R(M, M)$. We may call c and \bar{c} the first and second Casimir operators associated with M . They are of course dependent on μ_M .

THEOREM 8. *If $(R, S; M)$ holds, then an element $\omega \in Z_R$ is contained in $\mathfrak{I}'(M)$ if and only if there exists an S -endomorphism α of M such that $c(\alpha) = \zeta\omega$.*

Proof. Suppose α is an S -endomorphism of M such that $c(\alpha) = \zeta\omega$. Then $\beta = \alpha^+ \mu_M$ is an R -homomorphism of M into $I(M)$, and

$$\beta t_M = \alpha^+ \mu_M t_M = c(\alpha) = \zeta\omega.$$

Hence $\omega \in \mathfrak{I}^1(M)$ by Theorem 7.

On the other hand, $\omega \in \mathfrak{I}'(M)$ implies by Theorem 7 the existence of an R -homomorphism $\beta: M \rightarrow I(M)$ such that $\beta t_M = \zeta\omega$. Now $\alpha = \beta \mu_M^{-1} \epsilon_M$ is an S -endomorphism of M such that

$$c(\alpha) = \alpha^+ \mu_M t_M = (\beta \mu_M^{-1} \epsilon_M)^+ \mu_M t_M = \beta t_M = \zeta\omega.$$

The dual of Theorem 8 is obtained by replacing \mathfrak{I} by $\bar{\mathfrak{I}}$ and c by \bar{c} .

The condition $(R, S; M)$ may be strengthened by demanding that $c = \bar{c}$. Let us denote the resulting condition by $(R, S; M)^+$. If R is a self-dual S -ring in the sense of (2), $(R, S; M)^+$ holds for all R -modules M (2). Comparing Theorem 8 and its dual we have

COROLLARY 1. *If $(R, S; M)^+$ holds then $\mathfrak{I}^1(M) = \bar{\mathfrak{I}}^1(M)$.*

We can also prove

COROLLARY 2. *If $(R, S; K(M))^+$ holds then $\mathfrak{I}^2(M) \cap S_x = \mathfrak{I}^1(M) \cap S_x$, while dually, if $(R, S; L(M))^+$ holds then $\bar{\mathfrak{I}}^2(M) \cap S_x = \bar{\mathfrak{I}}^1(M) \cap S_x$.*

Proof. By Corollary 1, $\mathfrak{I}^2(M) = \mathfrak{I}^1(K(M)) = \bar{\mathfrak{I}}^1(K(M))$. By Theorem 1 there exists a $Z_R \cap S_x$ -isomorphism of $H^1(K(M), M)$ onto $H^1(M, K(M))$. It follows that $\bar{\mathfrak{I}}^1(K(M)) \cap S_x \subseteq \mathfrak{I}^1(M, K(M)) = \mathfrak{I}^1(M)$. Since $\mathfrak{I}^1(M) \cap S_x \subseteq \mathfrak{I}^2(M)$, the corollary is proved.

The above results may be extended by using the recursion relations $\mathfrak{I}^{i+\alpha}(M) = \mathfrak{I}^i(K^\alpha(M))$ and $\bar{\mathfrak{I}}^{i+\alpha}(M) = \bar{\mathfrak{I}}^i(L^\alpha(M))$. Thus, for example, we have by Corollary 2 that if $(R, S; M)^+$ holds for all R -modules M , then for $i > 0$,

$$\mathfrak{I}^i(M) \cap S_x = \mathfrak{I}^1(M) \cap S_x$$

and

$$\bar{\mathfrak{I}}^i(M) \cap S_x = \bar{\mathfrak{I}}^1(M) \cap S_x.$$

The methods of §4 may be used to give further information concerning these

ideals in the case considered there. For example, if M is an $A' \otimes_U A$ -module we find that

$$\mathfrak{I}_{(C,D)}^i(A) \cap D\eta \subseteq \mathfrak{I}_{(C,D)}^i(M)$$

where the notation is that of §4.

To see how ideals of the kind considered here occur in the study of orders in algebras, see (3) and (4).

REFERENCES

1. H. Cartan and S. Eilenberg, *Homological algebra* (Princeton, 1956).
2. D. G. Higman, *Induced and produced modules*, Can. J. Math., 7 (1955), 490-508.
3. ———, *On orders in separable algebras*, Can. J. Math., 7 (1955), 509-515.
4. ———, *On representations of orders* (in preparation).
5. G. Hochschild, *On the cohomology theory for associative algebras*, Ann. Math., 46 (1945), 58-67.
6. ———, *On the cohomology groups of an associative algebra*, Ann. Math., 47 (1946), 568-579.
7. ———, *Cohomology and representations of algebras*, Duke Math. J., 14 (1947), 921-948.
8. ———, *Relative Homological algebra*, Amer. Math. Soc. Trans., 82 (1956), 246-269.

Montana State University and
University of Michigan

SOME REMARKS ON NOETHERIAN RINGS

MICHIO YOSHIDA

In his lecture at the University of Kyoto on September 23, 1955, Professor Artin gave an important theorem on Noetherian rings, which seems to have not a few interesting consequences. It is the purpose of our present note to point out one of them. We begin by quoting a special case of the theorem.

THEOREM. *Let R be a Noetherian ring with unit element, and a, b ideals of R . Then there exists a positive integer d such that*

$$a^n \cap b = a^{n-r}(a^r \cap b) \quad n \geq r \geq d.$$

Proof. Let $\{a_1, \dots, a_m\}$ be a system of generators of a , and consider the polynomial ring $R[x] = R[x_1, \dots, x_m]$. Denote by A_r the set of forms of degree r in $R[x]$, and by B_r the set of all the forms $f(x)$ of degree r such that $f(a_1, \dots, a_m) \in b$. A_r is a R -module, B_r is a submodule of A_r , and obviously $A_{n-r} \cdot B_r \subseteq B_n$ for $n \geq r$. We select a finite system of forms $f_i(x)$, $1 \leq i \leq l$, from $\{B_r; r = 0, 1, 2, \dots\}$ such that any form $f(x)$ of $\{B_r; r = 0, 1, 2, \dots\}$ may be represented as

$$f = \sum_{i=1}^l \phi_i \cdot f_i,$$

where ϕ_i 's are forms of $R[x]$. Denote by d the maximum of the degrees of $f_i(x)$, $1 \leq i \leq l$, then for $n \geq r \geq d$, $A_{n-r} \cdot B_r = B_n$, namely $a^n \cap b = a^{n-r}(a^r \cap b)$.

By taking a principal ideal for b , we obtain the following:

COROLLARY. *Let a be an ideal of R , and a a nonzero-divisor of R , then there exists a positive integer d such that*

$$a^n : Ra = a^{n-r}(a^r : Ra) \quad n \geq r \geq d,$$

consequently

$$a^n : Ra \subseteq a^{n-r}.$$

Though Professor Artin did not mention this corollary, the last formula $a^n : Ra \subseteq a^{n-r}$ is of some interest. This is really a satisfactory generalization of a well-known theorem (1, p. 699, Lemma 9; 5, p. 38, Lemma 1). We would refer readers to a remark by Samuel on this kind of formula (2, p. 34). This formula enables us to sharpen one of his results (2, p. 23) as follows.

Received October 14, 1955.

THEOREM 1. *Let \mathfrak{a} be an ideal of Noetherian ring R . If \mathfrak{a} contains at least one nonzero-divisor, then there exists an element a of \mathfrak{a} such that*

$$\mathfrak{a}^{n+r}: Ra = \mathfrak{a}^n$$

for sufficiently large n , where r is determined by $a \in \mathfrak{a}^r$ and $a \notin \mathfrak{a}^{r+1}$.

Proof. Put

$$\mathfrak{n} = \bigcap_{n=1}^{\infty} \mathfrak{a}^n, \quad {}^*R = R/\mathfrak{n}, \quad {}^*a = a/\mathfrak{n}.$$

It is easily seen e.g. by the intersection theorem (4, p. 180, Theorem 3) that ${}^*\mathfrak{a}$ contains at least one nonzero-divisor and that any prime ideal of the zero ideal of *R is closed and not open in ${}^*\mathfrak{a}$ -adic topology. So Samuel's observations on the ring of forms $F({}^*\mathfrak{a}) = \sum {}^*\mathfrak{a}^i / {}^*\mathfrak{a}^{i+1}$ (2, p. 22-23) ensure the existence of a superficial element *a of some degree r with respect to ${}^*\mathfrak{a}$, which is not a zero-divisor. Hence ${}^*\mathfrak{a}^{n+r}: {}^*R {}^*a = {}^*\mathfrak{a}^n$ for sufficiently large n . Any element in the residue class *a will have the property required in the theorem.

COROLLARY. *Under the same assumption on \mathfrak{a} , there exist positive integers r, n_0 such that*

$$\mathfrak{a}^{nr+r}: \mathfrak{a}^{mr} = \mathfrak{a}^{nr}, \quad n > n_0.$$

We do not know whether we can always take 1 for r in this corollary, but Samuel (3, p. 177, Theorem 10) tells us the following:

THEOREM. *Let A be a local ring with the maximal ideal \mathfrak{m} , and let \mathfrak{q} be an \mathfrak{m} -primary ideal. Suppose \mathfrak{m} contains at least one nonzero-divisor, then*

$$\mathfrak{q}^n: \mathfrak{q} = \mathfrak{q}^{n-1} \quad \text{for sufficiently large } n.$$

Proof. In the case that the residue field $k = A/\mathfrak{m}$ is infinite, his assertion is substantiated by the existence of a superficial element of degree 1 with respect to \mathfrak{q} , which is not a zero-divisor (2, p. 23). The other case that k is finite shall be reduced to the former case by the following device. Form the polynomial ring $A[x]$ in an indeterminate X , then form the ring of quotients A^* of $\mathfrak{m}A[x]$ with respect to $A[x]$. The residue field of A^* is $k(x)$, hence

$$\mathfrak{q}^n A^*: \mathfrak{q} A^* = \mathfrak{q}^{n-1} A^*.$$

Notice that

$$(\mathfrak{q}^n A^*: \mathfrak{q} A^*) \cap A = \mathfrak{q}^n: \mathfrak{q}, \quad \mathfrak{q}^{n-1} A^* \cap A = \mathfrak{q}^{n-1}.$$

Before we transform the above theorems by "globalization," we shall recall some definitions and well-known facts. Let \mathfrak{z} be a prime ideal of R , and \mathfrak{q} a \mathfrak{z} -primary ideal. The \mathfrak{z} -primary component of \mathfrak{q}^n is called n th symbolic power of \mathfrak{q} , and usually denoted by $\mathfrak{q}^{(n)}$. Let \mathfrak{a} be an ideal of R , and z_1, \dots, z_t be the minimal prime ideals of \mathfrak{a} . The intersection of the z_i -primary compo-

nents ($1 \leq i \leq l$) of \mathfrak{a}^n is called n th symbolic power of \mathfrak{a} , and denoted by $\mathfrak{a}^{(n)}$. If \mathfrak{q}_i denotes the \mathfrak{z}_i -primary component of \mathfrak{a} , then as is well known

$$\mathfrak{a}^{(n)} = \mathfrak{q}_1^{(n)} \cap \dots \cap \mathfrak{q}_l^{(n)}.$$

We denote by S the complement of

$$\bigcup_{i=1}^l \mathfrak{z}_i$$

in R , and form the ring of quotients R_s of S with respect to R in the Chevalley-Uzkov sense. We have then $\mathfrak{a}^{(n)} = \mathfrak{a}^n R_s \cap R$. Let

$$(0) = \mathfrak{q}_1^* \cap \dots \cap \mathfrak{q}_l^*$$

be a primary decomposition of the zero ideal of R , and let \mathfrak{z}_i^* be the prime ideal of \mathfrak{q}_i^* . Assume $\mathfrak{z}_i^* \cap S = \emptyset$ for $i = 1, \dots, s$ and $\mathfrak{z}_i^* \cap S \neq \emptyset$ for $i = s+1, \dots, l$. Then $\mathfrak{n} = \mathfrak{q}_1^* \cap \dots \cap \mathfrak{q}_s^*$ is the kernel of the canonical homomorphism of R into R_s . Contracting of ideals of R_s on R and extending of ideals of R to R_s both give one-to-one mappings between the set of all ideals of R_s and the set of ideals of R whose prime ideals are disjoint with S . These mappings are the inverse of each other and they are isomorphisms with respect to the ideal operations (\cap) and $(:)$. We are now in a position to verify the following:

THEOREM 2. *Let \mathfrak{a} be an ideal of a Noetherian ring R . Suppose that any minimal prime ideal of \mathfrak{a} is not a prime ideal of (0) . Then there exist an element a of \mathfrak{a} and a positive integer n_0 such that*

$$\mathfrak{a}^{(n+r)} : Ra = \mathfrak{a}^{(n)}, \quad n \geq n_0$$

where r satisfies $a \in \mathfrak{a}^{(r)}$ and $a \notin \mathfrak{a}^{(r+1)}$. Moreover

$$\mathfrak{a}^{(n+m)} : \mathfrak{a}^{(m)} = \mathfrak{a}^{(n)}$$

for sufficiently large n and arbitrary $m \geq 0$.

REFERENCES

1. C. Chevalley, *On the theory of local rings*, Ann. Math. 44 (1943), 690-708.
2. P. Samuel, *Algèbre locale*, Mem. Sci. Math., No. 123 (Paris, 1953).
3. ———, *Sur la notion de multiplicité en algèbre et en géométrie algébrique*, J. Math. pur. et appl. 30 (1950), 159-205.
4. O. Zariski, *Generalized semi-local rings*, Sum. Bras. Math. 1 (1946), 169-195.
5. ———, *Theory and applications of holomorphic functions on algebraic varieties over arbitrary fields*, Mem. Amer. Math. Soc., No. 5 (New York, 1951).

Hiroshima University

SPACES OF DIMENSION ZERO

BERNHARD BANASCHEWSKI

1. Introduction. In a recent paper (1) it was remarked that the theory of zero-dimensional spaces is exactly that part of general topology which can be described in terms of equivalence relations. Here, it will be shown how this idea can be used to obtain the following characterizations of certain types of zero-dimensional spaces:

Any compact zero-dimensional space which has a denumerable basis for its open sets and is dense in itself is homeomorphic to the space of 2-adic integers.

Any locally compact zero-dimensional space which is non-compact, has a denumerable basis for its open sets and is dense in itself, is homeomorphic to the space of 2-adic numbers.

The first of these statements is a well-known theorem (6; vol. 2, §40, II), whilst the second one, here occurring as a simple consequence of the former, was proved in (5). However, in both cases, the methods employed are rather different from ours which are, in fact, no more than a refinement of arguments used in (1). In similar ways, the following assertions concerning non-archimedean metric spaces will be proved:

The number of inequivalent non-archimedean metrics on a non-compact zero-dimensional space which has a denumerable basis for its open sets and is dense in itself, is at least equal to the power of the continuum.

Any separable non-archimedean metric space can be mapped by a metric equivalence into the space of all formal power series with integral coefficients, taken with its so-called topology of formal convergence.

Any two n -adic metric spaces are metrically equivalent to each other.

2. Preliminaries. Topological terms, unless otherwise stated, will be used in the sense of (3). The term "space" will always be taken to mean "Hausdorff space which has a denumerable basis for its open sets." Zero-dimensionality means the existence of a basis for the open sets consisting of open-closed sets. Two metric spaces E and F are called metrically equivalent if there exists a homeomorphism from E onto F which is uniformly continuous in both directions. Two metrics on the same space E are called equivalent if the identity mapping of E onto itself is a metric equivalence with respect to these metrics. The topology of formal convergence (a term due to E. Witt) in the ring of all formal power series

$$p = \sum_{n \geq 0} c_n x^n,$$

Received May 4, 1956.

c_n arbitrary elements from a given ring and z an indeterminate, is obtained by taking the ideals (z^k) as a system of neighbourhoods of $p = 0$. The space of n -adic numbers, defined by completing the rational numbers with respect to the ring topology given by the ideals (n^k) , taken in its natural metric, will be referred to as the n -adic metric space. A uniform structure of a space E is here a "système fondamentale des entourages" in the sense of (3; chap. 11), compatible with the topology of E .

Equivalence relations on a set will be denoted by $\alpha, \beta, \gamma, \dots$. All equivalence relations considered here will be relations on some topological space E . The α -class to which $x \in E$ belongs will be called $\alpha(x)$. The α for which $\alpha(x) = E$ is called the all-relation. The number of α -classes into which E decomposes will be denoted by $|\alpha|$ and called the index of α . If each α -class is an open-closed set in E , α will be called open-closed. The expression $\alpha < \beta$ (" α is finer than β ") means that each α -class is contained in some β -class. If $\alpha < \beta$ and each β -class contains the same number of α -classes, this number will be denoted by $(\beta: \alpha)$, called the index of α in β . Writing down $(\beta: \alpha)$ will always be meant to imply the existence of this number.

As an immediate consequence of (1, Satz 10), one has:

LEMMA 1. *If a compact zero-dimensional space E possesses a uniform structure consisting of a decreasing sequence $\alpha_1 > \alpha_2 > \dots$ of open-closed equivalence relations for which α_1 is the all-relation on E and each $(\alpha_{i-1} : \alpha_i)$ equals 2, then E is homeomorphic to the space of 2-adic integers.*

According to (1), this homeomorphism is given by the following method: The α_k -classes in each α_{k-1} -class are taken to be numbered, in a fixed manner, by 0 and 1; and for each $x \in E$, $c_k(x)$ is defined as the number $\alpha_{k+2}(x)$ in $\alpha_{k+1}(x)$. Then, E can be mapped by

$$x \rightarrow p(x) = \sum_{k \geq 0} c_k(x) z^k$$

into the ring \mathfrak{P} of all power series in an indeterminate z with coefficients 0 and 1 from the prime field of characteristic 2. This mapping is a homeomorphism of E onto \mathfrak{P} if \mathfrak{P} is taken with its topology of formal convergence. In this topology, however, \mathfrak{P} is homeomorphic to the space of 2-adic integers.

Lemma 1 can be strengthened slightly: One can replace the hypothesis $(\alpha_i : \alpha_{i+1}) = 2$ by the weaker condition

$$(\alpha_i : \alpha_{i+1}) = 2^{n_i},$$

with some natural numbers n_i , for in this case there are, for each i , sequences

$$\alpha_i = \beta_1 > \beta_2 > \dots > \beta_{n_i} = \alpha_{i+1}$$

between α_i and α_{i+1} satisfying $(\beta_j : \beta_{j+1}) = 2$.

A further result from (1) needed here is:

LEMMA 2. *A space E is zero-dimensional if and only if its open-closed equivalence relations form a uniform structure of E . A compact E is zero-*

dimensional if and only if it has a uniform structure consisting of a decreasing sequence of open-closed equivalence relations of finite index.

Of course, the second part of this statement would no longer be true if the condition, always implicitly assumed here, that E have a denumerable basis were not satisfied.

Finally, a metric $|x, y|$ on a set E is called non-archimedean, if it satisfies the condition $|x, y| \leq \max\{|x, z|, |z, y|\}$ for any x, y and z from E . It is well known that any non-archimedean metric space is zero-dimensional.

3. The compact case. In order to prove the first statement in §1 it is now sufficient to show that any compact E of dimension zero and dense in itself possesses a uniform structure consisting of open-closed α_i such that $\alpha_i \supset \alpha_{i+1}$ and $(\alpha_i: \alpha_{i+1})$ is always a power of 2.

Let α_i ($i = 1, 2, \dots$) be a decreasing sequence of relations on E as given by Lemma 2. Since E is dense in itself, any α_i -class must consist of more than just one point. Therefore, any fixed α_i -class contains an arbitrarily large number of α_k -classes for suitably large k . From this, it can be deduced that there is also a decreasing sequence of open-closed equivalence relations β_i such that $\beta_i \subset \alpha_i$, $\beta_i \supset \alpha_{n(i)}$ for some suitable $n(i)$, and $(\beta_i: \beta_{i+1})$ is a power of 2.

Suppose that the first k members $\beta_1 \supset \beta_2 \supset \dots \supset \beta_k$ of this new sequence have already been determined. Then, by assumption, $\beta_k \supset \alpha_{n(k)}$. If $(\beta_k: \alpha_{n(k)})$ is defined and a power of 2, one can take

$$\beta_{k+1} = \alpha_{n(k)}, \quad n(k+1) = n(k) + 1.$$

Otherwise, let m be the largest number of $\alpha_{n(k)}$ -classes contained in any β_k -class and $2^s \geq m$. In any of the finitely many β_k -classes B , let m_B be the number of $\alpha_{n(k)}$ -classes and C a fixed one of these. Now, for a sufficiently large l_B , C contains more than $2^s - m_B + 1$ α_{l_B} -classes. By forming, if necessary, unions of these, one can obtain a decomposition of C into exactly $2^s - m_B + 1$ open-closed sets. These, together with the $m_B - 1$ $\alpha_{n(k)}$ -classes in B other than C decompose B into 2^s open-closed sets, and taking this for each B , one has a decomposition of this kind for E . The corresponding relation β is open-closed, satisfies $\beta \subset \alpha_{k+1}$ because of $\beta \subset \alpha_{n(k)} \subset \beta_k \subset \alpha_k$ and also $\beta \supset \alpha_l$ for any l greater than all l_B . Hence, one can put $\beta_{k+1} = \beta$ and $n(k+1)$ equal to, say, the first number greater than the l_B .

This completes the proof, since it was assumed that α_1 is the all-relation and β_1 , therefore, can be taken as α_1 . That the sequence β_i forms a uniform structure of E is, of course, an immediate consequence of $\beta_i \subset \alpha_i$.

As a corollary one has: *Any totally bounded non-archimedean metric space which is dense in itself is metrically equivalent to a subspace of the metric space of 2-adic integers.* For a space E of this type has a zero-dimensional compact \bar{E} as its metric completion which is also dense in itself, therefore homeomorphic to the space of 2-adic integers and hence metrically equivalent to it with respect to its metric induced from E and the natural metric for the 2-adic

integers in the latter. Also, since the space of \mathfrak{p} -adic integers for any prime ideal \mathfrak{p} of any number field is compact, dense in itself and has a denumerable basis, one obtains as a further consequence: *For any \mathfrak{p} , the space of \mathfrak{p} -adic integers is homeomorphic to the space of 2-adic integers.*

The first of these corollaries can be regarded as a partial strengthening of a theorem by Urysohn (6; vol. I, §23) according to which any zero-dimensional space is homeomorphic to a subspace of Cantor's compact zero-dimensional space which is, of course, homeomorphic to the space of 2-adic integers.

4. The non-compact locally compact case. If E is zero-dimensional, and not compact but locally compact, then it is the union of denumerably many disjoint open-closed compact sets: Since E has a denumerable basis for its open sets, it also has such a basis \mathfrak{B} consisting of open-closed sets. If, then, for each $x \in E$, V_x is an open neighbourhood with compact closure and $B_x \in \mathfrak{B}$ such that $x \in B_x \subseteq V_x$, these B_x are compact open-closed and have E as their union. Furthermore, there are only denumerably many of them and, hence, they can be arranged in a sequence B_i , $i = 1, 2, \dots$. Now, the disjoint sets

$$B_k^* = B_k - \bigcup_{i < k} B_i, \quad k = 1, 2, \dots$$

still have E as their union and are open-closed compact.

The considered space E being dense in itself, each of these B_k^* , since it is open in E , must also be dense in itself and therefore homeomorphic to the space of 2-adic integers. Furthermore, as the B_k^* are open-closed, E is the topological sum in the sense of (3; chap. 1) of its compact subspaces B_k^* , hence homeomorphic to the sum of denumerably many spaces of 2-adic integers. Finally, as the space of the 2-adic numbers is itself a space of this type, E is homeomorphic to it.

In exactly the same manner as above one obtains the corollary: *For any \mathfrak{p} , the space of \mathfrak{p} -adic numbers is homeomorphic to the space of 2-adic numbers.*

5. The number of distinct non-archimedean metrics of a zero-dimensional space. Let E be the space in question. It possesses a sequence $\alpha_1 > \alpha_2 > \dots$ of open-closed equivalence relations where $\cap \alpha_i(x) = x$ and the $\alpha_i(x)$ form a neighbourhood basis for each $x \in E$, originating from one of its non-archimedean metrics which are known to exist (1). The connection between the α_i and the metric is such that

$$(*) \quad |x, y| = 2^{-i} \text{ if } x \alpha_i y, \quad i = 1, 2, \dots, k; \quad i \neq k+1.$$

defines an equivalent metric (1). If, now, each $|\alpha_i|$ is finite, E is totally bounded with respect to this metric. Therefore, under the hypothesis that E is not totally bounded in each of its non-archimedean metrics (the other case will be considered later on) one can assume that, for some i , E decomposes into infinitely many α_i -classes. Obviously, no generality is lost by taking $i = 1$.

Since E is separable, the open-closed α_i -classes are a denumerable collection of sets, say, C_1, C_2, \dots . Then, with respect to a given increasing sequence k_1, k_2, \dots of natural numbers, one can decompose C_i into its α_{k_i} -classes. The decomposition of E into open-closed sets thus obtained gives an open-closed equivalence relation β_1 which can be used to define the sequence $\beta_k = \beta_1 \wedge \alpha_k$ where \wedge denotes taking the lattice theoretic meet of two equivalence relations (2). In the manner given by the formula (*), the sequence $\beta_1 > \beta_2 > \dots$ defines a new non-archimedean metric on E . The number of metrics that can be obtained in this way is equal to the number of increasing sequences of natural integers, hence equal to the cardinal number c of the continuum. However, these c different metrics need not all be inequivalent to each other. In order to prove the assertion stated in §1 it will now be shown that this set of metrics splits into c different equivalence classes.

Let k^*_i and k_i be two different increasing sequences of natural integers and β^*_k, β_k the two corresponding sequences of open-closed equivalence relations. The metrics defined by β^*_k and β_k will be equivalent if and only if to each β_k there exists a $\beta^*_e < \beta_k$ and vice versa. In particular, one then has a relation of the type $\beta_1 > \beta^*_i > \beta_k$ with suitable i and k . Now, by definition of β_e , the β_e -classes contained in C_m will be equal to the β_1 -classes in C_m for all sufficiently large m : as $\beta_e = \beta_1 \wedge \alpha_e$, the β_e -classes on C_m are intersections of α_{k_m} -classes and α_e -classes. If m is large enough, one has $\alpha_{k_m} < \alpha_e$ anyway, so these intersections will merely be α_{k_m} -classes, and these are also the β_1 -classes in C_m . The relation $\beta_1 > \beta^*_i > \beta_k$ therefore implies that the β^*_i -classes in all C_m for sufficiently large m are also equal to the β_1 -classes, and this then gives the result: From a certain $m = m_0$ onwards β_1 and β^*_1 are equal in C_m .

Now, let the sequence $\alpha_1 > \alpha_2 > \dots$ satisfy this further condition: Any α_i -class decomposes into more than one α_{i+1} -class. Then, β_1 and β^*_1 can only be equal on C_m if $k_m = k^*_m$. In this case, therefore, one obtains that the two sequences k_i and k^*_i are equal from a suitable index onwards. Since to any increasing sequence of natural integers there exist only denumerably many other such sequences coinciding with it from a suitable index onwards, the c different metrics defined above group into equivalence classes of at most denumerably many metrics each; the number of these classes will then still be c .

The restriction just placed on the sequence $\alpha_1 > \alpha_2 > \dots$ can be shown to be satisfied if not by the α_i themselves, then at least by suitable modifications of them. The basis for this will be that each open-closed set in E , being infinite since E is dense in itself, can be decomposed into two open-closed sets. Using this, one may define new relations α^*_i in the following way: $\alpha^*_1 = \alpha_1$. If α^*_n is already defined such that $\alpha^*_n < \alpha_n$, decompose each $\alpha^*_n \wedge \alpha_{n+1}$ -class into two open-closed sets and define α^*_{n+1} by the resulting decomposition of E . The sequence $\alpha^*_1 > \alpha^*_2 > \dots$ has all desired properties and using it, if necessary, in place of the original α_i one obtains the existence of c inequivalent non-archimedean metrics on E .

The remaining case to be considered is that E is totally bounded in all its non-archimedean metrics. This property implies, as will be shown now, the compactness of E . Let $\alpha_1 \geq \alpha_2 \geq \dots$ be chosen as above and take any open-closed equivalence relation α on E . As the sequence α_i defines a non-archimedean metric on E by (*) so does the sequence $\alpha'_i = \alpha \wedge \alpha_i$. E being totally bounded in this metric, there are only finitely many α -classes. Hence, any decomposition of E into open-closed sets must be finite. From this it follows that any denumerable open-closed covering of E contains a finite covering, for if

$$\bigcup_{i=1}^{\infty} B_i = E,$$

B_i open-closed, then the

$$B_i^* = B_i - \bigcup_{k < i} B_k$$

give a decomposition of E into open-closed sets which will, of course, only be finite if $B_i^* = \emptyset$ and therefore

$$\bigcup_{k < i} B_k \supseteq B_i$$

from some $i = i_0$ onwards; this gives

$$\bigcup_{k < i_0} B_k = E.$$

Finally, as E is zero-dimensional, one concludes from this that any open covering of E , having a denumerable open-closed covering as a refinement, contains a finite covering. E therefore is compact.

In all, it is then proved that a zero-dimensional space which is dense in itself either has at least c inequivalent non-archimedean metrics or is compact, in which case, of course, all its metrics are equivalent.

6. Imbeddings by metric equivalences. Let P now be the ring of all formal power series

$$\sum_{n \geq 0} c_n x^n$$

with integral coefficients, taken with its topology of formal convergence. This space \mathfrak{P} is a universal space for the separable non-archimedean metric spaces in the sense that any such space can be mapped into \mathfrak{P} by a metric equivalence. The mapping which will do this can again be defined as follows (1):

Let $\alpha_1 \geq \alpha_2 \geq \dots$ be a sequence of open-closed equivalence relations on E resulting from its given metric. Then, the α_i -classes in the different α_{i-1} -classes can be regarded as numbered in a fixed manner. With respect to this numbering, let $c_n(x)$ be the number of $\alpha_n(x)$ in $\alpha_{n-1}(x)$ for $x \in E$; then, in exactly the same way as in §1 in the special case of the compact spaces (Lemma 1) the mapping

$$p: x \rightarrow p(x) = \sum c_n(x)z^n$$

is a metric equivalence of E into \mathfrak{B} . Of course, in special cases, p may even be a metric isomorphism, but since the sequence $\alpha_1 > \alpha_2 > \dots$ determines the metric only up to equivalence, this will not be so in general.

\mathfrak{B} is obviously itself a separable non-archimedean metric space, since the denumerable set of all integral polynomials is dense in \mathfrak{B} . Therefore \mathfrak{B} is, in a sense, a minimal universal space for this type of space and, of course, characterized by this property.

For a more restricted class of metric spaces than the one just considered, one can obtain a universal space of an even simpler nature than the space \mathfrak{B} . A non-archimedean metric on a separable space E will be called *evenly locally compact* if it can be represented — up to equivalence — by a decreasing sequence $\alpha_1 > \alpha_2 > \dots$ of open-closed equivalence relations such that $\alpha_1(x)$ is compact for each $x \in E$ and the (necessarily finite) indices $(\alpha_{n-1}: \alpha_n)$ exist (see §2). Then, the following holds: *Each separable space with an evenly locally compact non-archimedean metric is metrically equivalent to a subspace of the metric space of 2-adic numbers.* The class of spaces admitted here includes, of course, all the n -adic metric spaces for any natural integer n and, more generally, the spaces of all separable locally compact groups whose topologies are given by a denumerable decreasing sequence of invariant subgroups as neighbourhood basis for the unit element.

By the preceding construction, a space E of the type now considered is metrically equivalent to a subset of \mathfrak{B} contained in the set given by all

$$\sum_{n \geq 0} a_n z^n, \quad 0 \leq a_n < j_n,$$

where j_0 denotes the (possibly infinite) number of α_1 -classes in E and j_n , $n \geq 1$, the index $(\alpha_n: \alpha_{n+1})$. Now, for any natural integer a one has

$$a = \varphi_0(a) + \varphi_1(a) 2 + \dots + \varphi_s(a) 2^s,$$

$\varphi_s(a)$ equal to 0 or 1, with some suitable s . In particular, then, any a , $0 \leq a < j_n$, can be written as

$$\varphi_0(a) + \varphi_1(a) 2 + \dots + \varphi_{s_n}(a) 2^{s_n} \text{ for } n \geq 1.$$

Using this, we can define a mapping φ of

$$w = \sum a_n z^n, \quad 0 \leq a_n < j_n, \quad n = 0, 1, 2, \dots,$$

by

$$\begin{aligned} \varphi(w) &= \varphi_{s_0}(a_0) z^{-s_0} + \varphi_{s_0-1}(a_0) z^{-s_0+1} + \dots + \varphi_1(a_0) z^{-1} + \varphi_0(a_0) \\ &\quad + \sum_{n=1}^m (\varphi_1(a_n) z + \varphi_2(a_n) z^2 + \dots + \varphi_{s_n}(a_n) z^{s_n}) z^{s_1+s_2+\dots+s_{n-1}}. \end{aligned}$$

$\varphi(w)$ is an element of the ring \mathfrak{L} of all Laurent forms

$$\sum_{n=0}^e a_n z^n$$

in z with integral coefficients, which, again, will be taken with the topology of formal convergence. Now, $w_1 \equiv w_2 \pmod{z^{e+1}}$ implies $a_n = b_n$, $n = 0, 1, \dots, e$, for the coefficients a_n of w_1 , and b_n of w_2 and therefore $\varphi_i(a_n) = \varphi_i(b_n)$ for $n = 0, 1, \dots, e$ and all corresponding i . From this it follows that

$$\varphi(w_1) \equiv \varphi(w_2) \pmod{z^{s_1+s_2+\dots+s_e+1}}.$$

Similarly, the converse holds, and since the

$$(z^{s_1+s_2+\dots+s_e+1})$$

define a metric equivalent to that defined by the (z^e) , φ is a metric equivalence. Furthermore, all $\varphi(w)$ lie in a part of \mathfrak{L} which is itself metrically equivalent to the 2-adic metric space. This proves the above assertion.

In the case of non-compact separable evenly locally compact non-archimedean metric spaces which are dense in themselves one can easily obtain a much stronger result: *Any such space is metrically equivalent to the metric space of 2-adic numbers.* A space E of this type is, of course, the sum of its compact open-closed α_1 -classes (see above) K_i , $i = 1, 2, \dots$, and each of these is homeomorphic to the space of 2-adic integers. If C_i , $i = 1, 2, \dots$, is the complete system of residue classes, with respect to addition, of all 2-adic numbers modulo the 2-adic integers, then K_i can be mapped homeomorphically onto C_i for each i . This gives a mapping ψ defined on E which is a metric equivalence. For, up to equivalence between metrics on E , the K_i can be taken to be the "unit spheres" in E , and this is what the C_i are in the 2-adic metric space. Then, since ψ carries unit spheres into unit spheres and is, in its restriction to these, a metric equivalence, it also is this for E as a whole.

In particular, this proves the last statement in §1 since the n -adic metric spaces are of the type just considered. More so, one can say that *there are only two essentially different, i.e., inequivalent, separable and evenly locally compact non-archimedean metric spaces which are dense in themselves*: The metric spaces of 2-adic integers and 2-adic numbers.

As a concluding remark, it may be pointed out that the metric product spaces of any finite number of p_i -adic metric spaces studied in (5) are all metrically equivalent to each other. This follows from the last remark and the fact that the metric product of p_i -adic metric spaces ($i = 1, 2, \dots, r$), is metrically equivalent to the $p_1 p_2 \dots p_r$ -adic metric space, which is an immediate consequence of (4; p. 96, Satz 5). There is a remarkable contrast between this and the results in (5) according to which an *isometric* mapping of one product of this type into another one can only exist if the product of the corresponding sets of prime numbers for the first space is less than or equal to that for the second space. This, then, goes to show that metric equivalence leads to much wider classes than the more restrictive concept of isometry.

REFERENCES

1. B. Banaschewski, *Ueber nulldimensionale Räume*, Math. Nachr. 13 (1955), 129-140.
2. G. Birkhoff, *Lattice theory*, Amer. Math. Soc. Coll. Publ. XXV (1948).
3. N. Bourbaki, *Topologie générale*, chaps. I and II, Act. sci. et ind. (Paris, 1948).
4. M. Deuring, *Algebren* (Berlin, 1935).
5. G. E. N. Fox, *The topology of p -adic number fields*, doctoral thesis, McGill University, 1955.
6. C. Kuratowski, *Topologie I and II*, Monografie Matematyczne, XX and XXI (Warsaw, 1948 and 1950).

*Hamilton College,
McMaster University,
Hamilton, Ontario*

MATRICES WITH ELEMENTS IN A BOOLEAN RING

A. T. BUTSON

1. Introduction. Let \mathfrak{B} be a Boolean ring of at least two elements containing a unit 1. Form the set \mathfrak{M} of matrices A, B, \dots of order n having entries a_{ij}, b_{ij}, \dots ($i, j = 1, 2, \dots, n$), which are members of \mathfrak{B} . A matrix U of \mathfrak{M} is called *unimodular* if there exists a matrix V of \mathfrak{M} such that $VU = I$, the identity matrix. Two matrices A and B are said to be *left-associates* if there exists a unimodular matrix U satisfying $UA = B$. The main results in this paper are the constructions of two canonical forms for left-associated matrices of \mathfrak{M} . The first form may be described very simply; however, it lacks the desirable property of containing the maximum possible number of rows which consist entirely of 0's. Although the second has this property, its description is quite complicated. They are somewhat similar to the well-known Hermite form for matrices with elements in a principal ideal ring (4); and, accordingly, use is made of them to establish analogues of several other familiar results concerning matrices with elements in a principal ideal ring. Although row equivalence (left-associativity) and a diagonal canonical form for equivalent matrices of \mathfrak{M} are mentioned in (2, pp. 164-165), the author has been unable to locate his results anywhere in the literature.

2. Properties of \mathfrak{B} . A Boolean ring may be defined as a ring whose elements are all idempotent. It is easily shown, see (2, pp. 154-155), that it is a commutative ring of characteristic two, in the usual sense. Then for any x in \mathfrak{B} , the element $x' = 1 + x$, called the complement of x , satisfies $x + x' = 1$, $xx' = 0$, and $(x')' = x$. Bell (1) observed that $x \vee y = x + x'y$ is the g.c.d. of x and y . Following is a summary of the less obvious but easily established properties of \mathfrak{B} which we shall use in the sequel:

$$(2.1) \quad xx = x;$$

$$(2.2) \quad xy = yx;$$

$$(2.3) \quad x + x = 0;$$

$$(2.4) \quad x + x' = 1, xx' = 0, (x')' = x;$$

$$(2.5) \quad \bigvee_{i=1}^n x_i = x_1 \vee x_2 \vee \dots \vee x_n = x_1 + x'_1x_2 + x'_1x'_2x_3 + \dots + x'_1x'_2 \dots x'_{n-1}x_n$$

is the g.c.d. of x_1, x_2, \dots, x_n ;

$$(2.6) \quad \left(\sum_{j=1}^t x_j \right) \left(\bigvee_{i=1}^n x_i \right) = \sum_{j=1}^t x_j, \quad t = 1, 2, \dots, n;$$

Received March 4, 1956.

$$(2.7) \quad \left(\sum_{i=1}^n x_i \right)' = x'_1 x'_2 \dots x'_n, \quad (x_1 x_2 \dots x_n)' = \sum_{i=1}^n x'_i;$$

$$(2.8) \quad \sum_{i=1}^n x_i = 0 \text{ if and only if } x_1 = x_2 = \dots = x_n = 0;$$

$$(2.9) \quad xy = 0 \text{ if and only if } xy' = x.$$

3. Canonical forms. In constructing the canonical forms only one type of *elementary operation* is needed, the addition to the elements of a row of x times the corresponding elements of another row, x being in \mathfrak{B} . Furthermore, this elementary operation can be accomplished by multiplying the given matrix on the left by an *elementary matrix*, namely the matrix obtained by performing the desired elementary operation upon the identity matrix I . If E is any elementary matrix, it follows from (2.3) that $EE = I$. Quite obviously then, any elementary matrix is unimodular, and a product of unimodular matrices is unimodular. To facilitate describing the constructions, we first establish a lemma.

LEMMA 3.1. For $0 \leq j \leq n$, let $A(j) = [B(j) H(n-j)]$ be the following matrix of \mathfrak{M} :

$$\begin{bmatrix} b_{11} & b_{12} & \dots & b_{1j} & 0 & 0 & \dots & 0 \\ b_{21} & b_{22} & \dots & b_{2j} & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ b_{j1} & b_{j2} & \dots & b_{jj} & 0 & 0 & \dots & 0 \\ b_{j+1,1} & b_{j+1,2} & \dots & b_{j+1,j} & h_{j+1,j+1} & 0 & \dots & 0 \\ b_{j+2,1} & b_{j+2,2} & \dots & b_{j+2,j} & h_{j+2,j+1} & h_{j+2,j+2} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ b_{n1} & b_{n2} & \dots & b_{nj} & h_{n,j+1} & h_{n,j+2} & \dots & h_{nn} \end{bmatrix}$$

where $A = A(n) = [B(n) H(0)]$, $H = A(0) = [B(0) H(n)]$, and $h_{pq} h_{qq} = 0$, $h_{pq} h_{pp} = h_{pq}$ for $p = q+1, q+2, \dots, n$; $q = j+1, j+2, \dots, n$. Then there exists a unimodular matrix U_j (which is a product of elementary matrices) such that multiplying $A(j)$ on the left by U_j leaves the last $n-j$ columns of $A(j)$ invariant, and replaces the elements b_{kj} of the j th column of $A(j)$ by elements h_{kj} , where $h_{kj} = 0$ for $k = 1, 2, \dots, j-1$; and $h_{kj} h_{jj} = 0$, $h_{kj} h_{kk} = h_{kj}$ for $k = j+1, j+2, \dots, n$. (In terms of matrices we have

$$A(j-1) = U_j A(j) = U_j [B(j) H(n-j)] = [B(j-1) H(n-j+1)],$$

where it is to be understood that although $H(n-j)$ is a submatrix of $H(n-j+1)$, $B(j-1)$ is not necessarily a submatrix of $B(j)$.

Let E_{kj} denote the elementary matrix obtained from I by adding x_{kj} times the elements of the k th row to the corresponding elements of the j th row, where $x_{1j} = b'_{1j}$:

$$x_{kj} = b'_{1j} b'_{1j} b'_{2j} \dots b'_{k-1,j}, \quad k = 2, 3, \dots, j-1;$$

$$x_{kj} = h'_{1k} b'_{1j} b'_{2j} \dots b'_{k-1,j}, \quad k = j+1, j+2, \dots, n.$$

It is quite obvious that adding x_{kj} times the elements of the k th row to the corresponding elements of the j th row, for $k = 1, 2, \dots, j-1$, does not affect the last $n-j$ columns of $A(j)$. For $q = j+1, j+2, \dots, n$; $k = j+1, j+2, \dots, n$;

$$\begin{aligned} x_{kj}h_{kq} &= h'_{kk}b'_{1j}b'_{2j} \dots b'_{k-1,j}h_{kq} = h_{kq}h'_{kk}b'_{1j}b'_{2j} \dots b'_{k-1,j} \\ &= h_{kq}h_{kk}h'_{kk}b'_{1j}b'_{2j} \dots b'_{k-1,j} = 0. \end{aligned}$$

Hence adding x_{kj} times the elements of the k th row to the corresponding elements of the j th row, for $k = j+1, j+2, \dots, n$, does not affect the last $n-j$ columns of $A(j)$ either. Then multiplying $A(j)$ on the left by the unimodular matrix

$$E_j = E_{nj}E_{n-1,j} \dots E_{j+1,j}E_{j-1,j} \dots E_{2j}E_{1j}$$

leaves the last $n-j$ columns unaltered, and replaces b_{jj} by

$$h_{jj} = b_{1j} \vee \dots \vee b_{jj} \vee (b_{j+1,j}h'_{j+1,j+1}) \vee \dots \vee (b_{nj}h'_{nn}).$$

Let F_{kj} , for $k = 1, 2, \dots, j-1, j+1, \dots, n$, denote the elementary matrix obtained from I by adding b_{kj} times the elements of the j th row to the corresponding elements of the k th row. Multiplication of $E_j A(j)$ on the left by F_{kj} obviously leaves the last $n-j$ columns invariant, and replaces b_{kj} by $h_{kj} = b_{kj} + b_{kj}h_{jj}$. By (2.6) and (2.3) $h_{kj} = b_{kj} + b_{kj} = 0$ for $k = 1, 2, \dots, j-1$. For $k = j+1, j+2, \dots, n$;

$$h_{kj}h_{jj} = (b_{kj} + b_{kj}h_{jj})h_{jj} = b_{kj}h_{jj} + b_{kj}h_{jj} = 0.$$

Using (2.7) we can write

$$\begin{aligned} h_{kj} &= (b_{kj} + b_{kj}h_{jj}) = b_{kj}h'_{jj} \\ &= b_{kj}(b_{kj}h'_{kk})'b'_{1j}b'_{2j} \dots b'_{jj}(b_{j+1,j}h'_{j+1,j+1})' \dots \\ &\quad (b_{k-1,j}h'_{k-1,k-1})'(b_{k+1,j}h'_{k+1,k+1})' \dots (b_{nj}h'_{nn})'; \end{aligned}$$

and since

$$\begin{aligned} b_{kj}(b_{kj}h'_{kk})' &= b_{kj}(1 + b_{kj}h'_{kk}) = b_{kj} + b_{kj}h'_{kk} \\ &= b_{kj}(1 + h'_{kk}) = b_{kj}h_{kk}, \end{aligned}$$

$$h_{kj} = b_{kj}h_{kk}b'_{1j}b'_{2j} \dots b'_{jj}(b_{j+1,j}h'_{j+1,j+1})' \dots$$

$$(b_{k-1,j}h'_{k-1,k-1})'(b_{k+1,j}h'_{k+1,k+1})' \dots (b_{nj}h'_{nn})',$$

and it is now obvious that $h_{kj}h_{kk} = h_{kj}$. Letting

$$F_j = F_{nj}F_{n-1,j} \dots F_{j+1,j}F_{j-1,j} \dots F_{2j}F_{1j},$$

it is apparent that $F_j E_j$ is the desired unimodular matrix U_j .

We remark that if $h_{jj} = 0$, then by (2.8)

$$b_{1j} = b_{2j} = \dots = b_{jj} = 0, \quad (b_{kj}h'_{kk}) = 0, \quad k = j+1, j+2, \dots, n,$$

whence by (2.9) $b_{kj}h_{kk} = b_{kj}$. So in this particular case we may choose $I'_j = I$.

We also note that if $h_{pp} = 0$, then the requirement that $h_{pq}h_{pp} = h_{pq}$ implies that $h_{pq} = 0$ for $q = j + 1, j + 2, \dots, p$.

THEOREM 3.1. *For any matrix A of \mathfrak{M} there exists a unimodular matrix U of \mathfrak{M} which is a product of elementary matrices and such that $UA = H$ has the following properties: $h_{pq} = 0$ for $q > p$, $h_{pq}h_{qq} = 0$, and $h_{pq}h_{pp} = h_{pq}$. (Note that if a diagonal element is 0, then the entire row consists of 0's). This form H is unique.*

Successive applications of Lemma 3.1 to $A = A(n)$ for $j = n, n - 1, \dots, 1$ yields $A(0) = [B(0)H(n)] = H$ and $U_1U_2 \dots U_n = U$ as the desired matrices.

To prove the uniqueness of H , let U and V be unimodular matrices such that $UA = H$ and $VA = G$ each have the form described above. (The result that a unimodular matrix has an inverse is implied by the succeeding corollary, which is established without assuming the uniqueness of H . To simplify matters, we use this now.) Then $U^{-1}H = A = V^{-1}G$, and $PH = G$, $QG = H$, where $P = VU^{-1}$ and $Q = UV^{-1}$. Thus, for fixed i , the following systems of equations must be satisfied:

$$\sum_{k=t}^n p_{ik}h_{kt} = g_{it}, \quad \sum_{k=t}^n q_{ik}g_{kt} = h_{it}, \quad t = 1, 2, \dots, n,$$

where $g_{it} = h_{it} = 0$ for $t > i$. Consider the first system. The last equation $p_{in}h_{nn} = 0$ and the condition $h_{ni}h_{nn} = h_{ni}$ imply $p_{in}h_{nn} = 0$ for $t = 1, 2, \dots, n$. Thus the first system is equivalent to

$$\sum_{k=t}^{n-1} p_{ik}h_{kt} = g_{it}, \quad t = 1, 2, \dots, n-1.$$

The last equation $p_{i,n-1}h_{n-1,n-1} = 0$ of this system and the condition $h_{n-1,t}h_{n-1,n-1} = h_{n-1,t}$ imply $p_{i,n-1}h_{n-1,t} = 0$ for $t = 1, 2, \dots, n-1$. Thus this system may be reduced to

$$\sum_{k=t}^{n-2} p_{ik}h_{kt} = g_{it}, \quad t = 1, 2, \dots, n-2.$$

Continuing this reduction for $t = n-2, n-3, \dots, i+1$ yields $p_{ik}h_{kt} = 0$ for $k > i$, $t = 1, 2, \dots, n$, and replaces the first system by the equivalent system

$$\sum_{k=t}^i p_{ik}h_{kt} = g_{it}, \quad t = 1, 2, \dots, i.$$

Similarly, the second system is equivalent to

$$\sum_{k=t}^i q_{ik}g_{kt} = h_{it}, \quad t = 1, 2, \dots, i.$$

Now $p_{it}h_{it} = g_{it}$ and $q_{it}g_{it} = h_{it}$ imply

$$\begin{aligned} g_{it} &= p_{it}h_{it} = p_{it}q_{it}g_{it} = p_{it}q_{it}p_{it}h_{it} \\ &= q_{it}p_{it}h_{it} = q_{it}g_{it} = h_{it}. \end{aligned}$$

Thus $p_{it}h_{it} = h_{it}$, and $p_{it}h_{it} = h_{it}$, for $t = 1, 2, \dots, i$, since $h_{it}h_{it} = h_{it}$. Similarly, $q_{it}g_{it} = g_{it}$ and $q_{it}g_{it} = g_{it}$. Now consider

$$\begin{aligned} p_{i,t-1}h_{i,t-1} + p_{it}h_{i,t-1} &= g_{i,t-1}, \\ q_{i,t-1}g_{i,t-1} + q_{it}g_{i,t-1} &= h_{i,t-1}. \end{aligned}$$

Multiplying by $h_{i,t-1}$ and $g_{i,t-1}$, respectively, gives

$$h_{i,t-1} = h_{i,t-1}g_{i,t-1}, \quad g_{i,t-1} = h_{i,t-1}g_{i,t-1},$$

since $h_{i,t-1}h_{i,t-1} = g_{i,t-1}g_{i,t-1} = 0$, and $p_{it}h_{i,t-1} = h_{i,t-1}$, $q_{it}g_{i,t-1} = g_{i,t-1}$.

Hence $g_{i,t-1} = h_{i,t-1}$, and also $p_{i,t-1}h_{i,t-1} = q_{i,t-1}g_{i,t-1} = 0$ which implies

$$p_{i,t-1}h_{i,t-1} = q_{i,t-1}g_{i,t-1} = 0, \quad t = 1, 2, \dots, i-1.$$

Next we consider

$$\begin{aligned} p_{i,t-2}h_{i,t-2} + p_{i,t-1}h_{i,t-2} + p_{it}h_{i,t-2} &= g_{i,t-2}, \\ q_{i,t-2}g_{i,t-2} + q_{i,t-1}g_{i,t-2} + q_{it}g_{i,t-2} &= h_{i,t-2} \end{aligned}$$

which are simply $p_{i,t-2}h_{i,t-2} + h_{i,t-2} = g_{i,t-2}$ and $q_{i,t-2}g_{i,t-2} + g_{i,t-2} = h_{i,t-2}$. Multiplying by $h_{i,t-2}$ and $g_{i,t-2}$, respectively, gives

$$h_{i,t-2} = g_{i,t-2}h_{i,t-2}, \quad g_{i,t-2} = h_{i,t-2}g_{i,t-2}.$$

Then $g_{i,t-2} = h_{i,t-2}$, and $p_{i,t-2}h_{i,t-2} = q_{i,t-2}g_{i,t-2} = 0$ which implies

$$p_{i,t-2}h_{i,t-2} = q_{i,t-2}g_{i,t-2} = 0, \quad t = 1, 2, \dots, i-2.$$

Continuing this procedure yields $g_{it} = h_{it}$ for $t = 1, 2, \dots, i$. Now letting i range from 1 to n establishes the identity of G and H . Hence H is unique.

COROLLARY 3.1. Every unimodular matrix of \mathfrak{M} is a product of a finite number of elementary matrices.

If the matrix A in the above theorem is unimodular, then $UA = H$, being a product of unimodular matrices, is also unimodular. Then there exists a matrix K such that $KH = I$. The properties of the elements of H are restrictive enough to require that $H = K = I$. Since U is a product of elementary matrices, say $E_i E_{i-1} \dots E_1$, we have $E_i E_{i-1} \dots E_1 A = I$. Hence $A = E_1 E_2 \dots E_i$, the desired result. We remark that it is now obvious that $AU = I$ so that $U = A^{-1}$.

The canonical form H does not have, in general, the maximum possible number of rows whose elements are all 0's that could be obtained by elementary row operations on A . The succeeding lemma makes this apparent. Our procedure now will be to obtain a second canonical form for A by performing elementary operations on H that will replace a row wherever possible by a row of 0's and alter the form of H as little as possible.

LEMMA 3.2. Let H be the matrix described in the preceding theorem and $h_{j_1}, h_{j_1 j_1}, \dots, h_{j_1 j_1}, j_1 < j_2 < \dots < j_i$, be the diagonal elements in the last

$n - j + 1$ columns of H that are different from 0. Then a necessary and sufficient condition that there exists a unimodular matrix V_j , such that multiplication of H on the left by V_j replaces h_{jj} by 0, and leaves invariant the last $n - j$ columns of H and any row which consists entirely of 0's, is that $h_{jj}h_{j1j_1}h_{j2j_2}\dots h_{jnj_n} = 0$.

The most general sequence of elementary operations that could be performed on the rows of H and leave the necessary things invariant is: the addition of an arbitrary multiple, say x_{jr} , of the elements of the j th row to the corresponding elements of j_r th row, for $r = 1, 2, \dots, l$; then the addition of say $y_{jr}h'_{j_rj_r}$, where y_{jr} is arbitrary, times the elements of the j_r th row to the corresponding elements of the j th row, for $r = 1, 2, \dots, l$. This replaces h_{jj} by

$$h_{jj} + \sum_{r=1}^l y_{jr}h'_{j_rj_r}(h_{jjs} + x_{jr}h_{js}),$$

which is simply

$$h_{jj} + h_{jj} \sum_{r=1}^l y_{jr}x_{jr}h'_{j_rj_r},$$

since

$$h'_{j_rj_r}h_{jrs} = 0.$$

In order to be able to replace h_{jj} by 0, under the required conditions it is then necessary that there exist

$$x_{jr} \text{ and } y_{jr}, \quad r = 1, 2, \dots, l,$$

such that

$$h_{jj} + h_{jj} \sum_{r=1}^l y_{jr}x_{jr}h'_{j_rj_r} = 0.$$

By adding h_{jj} to both sides we obtain the equivalent condition

$$h_{jj} \sum_{r=1}^l y_{jr}x_{jr}h'_{j_rj_r} = h_{jj}.$$

Since

$$y_{jr}x_{jr}h'_{j_rj_r} \bigvee_{s=1}^l h'_{j_sj_s} = y_{jr}x_{jr}h'_{j_rj_r}$$

by (2.6), we have

$$\begin{aligned} h_{jj} \bigvee_{s=1}^l h'_{j_sj_s} &= \left(h_{jj} \sum_{r=1}^l y_{jr}x_{jr}h'_{j_rj_r} \right) \bigvee_{s=1}^l h'_{j_sj_s} \\ &= h_{jj} \sum_{r=1}^l \left(y_{jr}x_{jr}h'_{j_rj_r} \bigvee_{s=1}^l h'_{j_sj_s} \right) \\ &= h_{jj} \sum_{r=1}^l y_{jr}x_{jr}h'_{j_rj_r} \\ &= h_{jj}. \end{aligned}$$

But this last relation implies, by (2.9), (2.7), and (2.4), that

$$h_{jj}h_{j_1j_1}h_{j_2j_2}\dots h_{j_lj_l} = 0.$$

Hence the condition is necessary.

Conversely, suppose

$$h_{jj}h_{j_1j_1}\dots h_{j_lj_l} = 0.$$

Then

$$h_{jj} \bigvee_{s=1}^t h'_{j_sj_s} = h_{jj},$$

and the aforementioned sequence of operations with

$$x_{jr} = 1, \quad (r = 1, 2, \dots, t),$$

$$y_{j_1} = x, \quad y_{jr} = h_{j_1j_1}h_{j_2j_2}\dots h_{j_{r-1}j_{r-1}} \quad (r = 2, 3, \dots, t),$$

replaces h_{jj} by

$$h_{jj} + h_{jj} \sum_{r=1}^t y_{jr}x_{jr}h'_{jrjr} = h_{jj} + h_{jj} \bigvee_{r=1}^t h'_{jrjr} = h_{jj} + h_{jj} = 0$$

and leaves the necessary things invariant. Thus the condition is sufficient and the lemma is proved.

Let us now determine precisely what happens to the elements in the first j columns of H when h_{jj} is replaced by 0 in the manner described in Lemma 3.2. Since

$$h'_{jrjr}h_{jrq} = 0,$$

h_{jq} for $q = 1, 2, \dots, j-1$, is replaced by

$$h_{jq} + h_{jq} \bigvee_{r=1}^t h'_{jrjr}.$$

It is necessary that

$$h_{jj}h_{j_1j_1}\dots h_{j_lj_l} = 0,$$

so that

$$h_{jj} \bigvee_{r=1}^t h'_{jrjr} = h_{jj}.$$

Using this and the fact that $h_{jq} = h_{jq}h_{jj}$, we have

$$\begin{aligned} h_{jq} + h_{jq} \bigvee_{r=1}^t h'_{jrjr} &= h_{jq} + h_{jq}h_{jj} \bigvee_{r=1}^t h'_{jrjr} \\ &= h_{jq} + h_{jq}h_{jj} = h_{jq} + h_{jq} \\ &= 0. \end{aligned}$$

Thus replacing h_{jj} by 0 replaces h_{jq} , for $q = 1, 2, \dots, j-1$, by 0 also. For $r = 1, 2, \dots, t$ and $q = 1, 2, \dots, j$, h_{jrq} is replaced by

$$d_{jrq} = h_{jq} + h_{jrq}.$$

We observe that

$$d_{jrq}h_{qq} = (h_{jq} + h_{jr,q})h_{qq} = h_{jq}h_{qq} + h_{jr,q}h_{qq} = 0,$$

so that the property of H that the product of an element with the diagonal element above it be 0 is preserved. Although

$$d_{jrq}h_{jr,jr} \neq d_{jr,q}$$

in general, we note that

$$\begin{aligned} d_{jrq}(h_{jj} \vee h_{jr,jr}) &= (h_{jq} + h_{jr,q})(h_{jj} \vee h_{jr,jr}) \\ &= h_{jq}(h_{jj} \vee h_{jr,jr}) + h_{jr,q}(h_{jj} \vee h_{jr,jr}) \\ &= h_{jq}h_{jj}(h_{jj} \vee h_{jr,jr}) + h_{jr,q}h_{jr,jr}(h_{jj} \vee h_{jr,jr}) \\ &= h_{jq}h_{jj} + h_{jr,q}h_{jr,jr} = h_{jq} + h_{jr,q} \\ &= d_{jr,q}. \end{aligned}$$

We also see that, for $q = 1, 2, \dots, j$,

$$h_{jq}h_{j_1j_1} \dots h_{j_1j_1} = h_{jq}h_{jj}h_{j_1j_1} \dots h_{j_1j_1} = 0.$$

Hence $h_{jr,q}$ is replaced by an element

$$d_{jr,q} = h_{jq} + h_{jr,q}$$

such that

$$d_{jrq}h_{qq} = 0, \quad h_{jq}h_{jj} = h_{jq}, \quad h_{jq}h_{j_1j_1} \dots h_{j_1j_1} = 0,$$

$$d_{jrq}(h_{jj} \vee h_{jr,jr}) = d_{jr,q}.$$

Now let H_1 denote the matrix resulting from replacing h_{jj} by 0 according to the procedure just described. We want to consider the problem of replacing a diagonal element, say h_{ii} , of H_1 by 0 using elementary operations that leave invariant the last $n - i$ columns and any row whose elements are all 0's. Let

$$h_{i_1i_1}, h_{i_2i_2}, \dots, h_{i_ei_e}, \quad i < i_1 < i_2 < \dots < i_e < j,$$

denote the diagonal elements of H_1 between

$$h_{ii} \text{ and } h_{j_1j_1}$$

which are not 0. When we attempt to parallel the discussion of Lemma 3.2 we find that, although we can add

$$y_{ir}h'_{ir,i_r}, \text{ where } y_{ir} \text{ is arbitrary,}$$

times the elements of the i_r th row to the corresponding elements of the i th row, we can't add

$$y_{jr}h'_{jr,j_r}, \text{ where } y_{jr} \text{ is arbitrary,}$$

times the elements of the j_r th row to the corresponding elements of the i th row. In order to leave invariant the last $n - i$ columns of H_1 we must add instead

$y_{jr}h'_{jj}h_{jrjr}$, where y_{jr} is arbitrary,

times the elements of the j ,th row to the corresponding elements of the i th row. With only this change, however, we obtain the following result. A necessary and sufficient condition that h_{ii} can be replaced by 0, by means of elementary operations that leave invariant the last $n - i$ columns and any row consisting of 0's, is that

$$h_{ii}h_{i_1i_1} \dots h_{i_v i_v} (h_{j_1j_1} \vee h_{jj}) (h_{j_2j_2} \vee h_{jj}) \dots (h_{j_tj_t} \vee h_{jj}) = 0.$$

If this condition is satisfied, to replace h_{ii} by 0 we choose

$$x_{ir} = x_{js} = 1, \quad r = 1, 2, \dots, v, s = 1, 2, \dots, t.$$

That is, we first add the i th row to each succeeding row which does not consist entirely of 0's. Then a multiple of the elements of each of these rows is added to the corresponding elements of the i th row. Choosing the y 's appropriately, this replaces $h_{iq} (q = 1, 2, \dots, i)$, by

$$h_{iq} + h_{iq} \left\{ \bigvee_{r=1}^v h'_{irir} \vee \bigvee_{s=1}^t (h'_{jrjr} h'_{jj}) \right\} = h_{iq} + h_{iq} = 0.$$

We note that

$$h_{irq}, \quad q = 1, 2, \dots, i, r = 1, 2, \dots, v,$$

is replaced by

$$h_{irq} + h_{iq};$$

and

$$d_{jq}, \quad q = 1, 2, \dots, i, s = 1, 2, \dots, t,$$

is replaced by

$$d_{jq} + h_{iq} = h_{jq} + h_{jq} + h_{iq}.$$

Denote this matrix by H_2 .

We are now able to describe the procedure for obtaining from the first canonical form H the second canonical form, which we shall call C . Consider successively the products

$$h_{j_1j_1}h_{j_2j_2} \dots h_{j_{t-1}j_{t-1}}, \dots, h_{j_1j_1} \dots h_{jj}, h_{j_1j_1} \dots h_{jj}h_{i_1i_1} \dots$$

of the diagonal elements of H which are different from 0. If none of these are 0, then $C = H$. Otherwise, there is a first one, say

$$h_{jj}h_{j_1j_1} \dots h_{j_1j_1},$$

which is 0. In this case replace h_{jj} by 0 according to the procedure described in Lemma 3.2. Let

$$Z_j = (h_{j_1j_1} \vee h_{jj})(h_{j_2j_2} \vee h_{jj}) \dots (h_{j_tj_t} \vee h_{jj}),$$

and consider successively the products

$$Z_jh_{i_1i_1}, \dots, Z_jh_{i_v i_v}, \dots, h_{ii}, \dots,$$

If all of these are 0, then $C = H_1$. Otherwise, there is a first one, say

$$h_{i_1 i_1} \dots h_{i_q i_q} Z_{j_1}$$

which is 0. In this case, replace h_{ii} by 0 as before. Let

$$Z_{i,j} = (h_{i_1 i_1} \vee h_{ii}) \dots (h_{i_q i_q} \vee h_{ii}) (h_{j_1 j_1} \vee h_{jj} \vee h_{ii}) \dots (h_{j_l j_l} \vee h_{jj} \vee h_{ii}),$$

and let

$$h_{k_1 k_1}, \dots, h_{k_w k_w}, k_1 < k_2 < \dots < k_w,$$

denote the diagonal elements in the first $i - 1$ columns of H_2 which are not 0. Consider successively the products

$$Z_{i,j} h_{k_w k_w}, \dots, Z_{i,j} h_{k_2 k_2} \dots h_{k_1 k_1}.$$

If all of these are 0, then $C = H_2$. Otherwise, there is a first one, say

$$Z_{i,j} h_{k_w k_w} \dots h_{k_f k_f},$$

which is different from 0. Replace $h_{k_f k_f}$ by 0 in, what should be by now, the obvious manner. Continuing this procedure yields the desired matrix C . Obviously each one of these steps can be accomplished by multiplying the particular H_i on the left by a unimodular matrix V_i . Hence there is a unimodular matrix V such that $VH = C$.

Note that replacing h_{ii} by 0 affects the element in the (p, q) -position of H_i only if $q < i < p$. Then, for $c_{pp} \neq 0$, if we let

$$c_{q_1 q_1}, c_{q_2 q_2}, \dots, c_{q_{l_p} q_{l_p}}$$

denote the diagonal elements of C between $c_{q-1, q-1}$ and c_{pp} which are 0, we see that

$$c_{pq} = h_{pq} + h_{q_1 q} + h_{q_2 q} + \dots + h_{q_{l_p} q}.$$

(Although the c_{q, q_r} 's include any diagonal element that was originally 0 in H , say

$$h_{q_e q_e} = c_{q_e q_e},$$

this does not affect the representation of c_{pq} since the corresponding $h_{q_e q} = 0$.) We summarize all this in the following theorem.

THEOREM 3.2. *Let A be any matrix of \mathfrak{M} and $UA = H$ its first canonical form. Then there exists a unimodular matrix V such that $WA = VUA = VH = C$, where $W = VU$, has the following form: $c_{pq} = 0$ for $q > p$; if $c_{pp} = 0$, then $c_{pq} = 0$ for $q = 1, 2, \dots, n$; $c_{pq} c_{qq} = 0$; if $c_{pp} \neq 0$ and*

$$c_{q_1 q_1}, c_{q_2 q_2}, \dots, c_{q_{l_p} q_{l_p}}$$

denote the diagonal elements of C between $c_{q-1, q-1}$ and c_{pp} which are 0, then

$$c_{pq} = h_{pq} + h_{q_1 q} + h_{q_2 q} + \dots + h_{q_{l_p} q}.$$

Furthermore, it is impossible to replace a diagonal element of C by 0 using elementary operations that leave invariant the succeeding columns and any row which consists entirely of 0's. This form C is unique.

The proof that C is unique proceeds along the same lines as the proof of the uniqueness of H , and will be omitted. We wish to emphasize, however, that to ensure uniqueness it is absolutely necessary to add the elements of the k th row to the corresponding elements of each succeeding row whose elements are not all 0's, as the first step in replacing any diagonal element h_{kk} by 0.

THEOREM 3.3. *A necessary and sufficient condition that two matrices A and B of \mathfrak{M} be left-associates is that they have the same canonical form H (or C).*

If $PB = A$, where P is unimodular, let U be a unimodular matrix such that $UA = H$ is the first canonical form of A . Then $H = UPB = VB$, where $V = UP$ is unimodular, so that H is the first canonical form of B also. Conversely, suppose that E and F are unimodular matrices such that $EA = FB = H$ is the first canonical form of A and of B . Then $QB = A$, where $Q = E^{-1}F$ is unimodular, and A and B are left-associates.

4. Mutual left-divisibility, g.c.r.d., and l.c.l.m. Two matrices A and B of \mathfrak{M} are said to be *mutually left-divisible* if and only if there exist matrices R and T of \mathfrak{M} such that $RA = B$ and $TB = A$. It is well known that the concepts of mutual left-divisibility and left-associativity are equivalent for matrices with elements in a principal ideal ring. Steinitz (5) has shown their equivalence for matrices with elements in an algebraic domain. Kaplansky (3) considered this problem and obtained some results based on the radical of a ring. We now show that the two concepts are equivalent for matrices of \mathfrak{M} . If A and B are left-associates so that $PA = B$, where P is unimodular, then $P^{-1}B = A$ and A and B are mutually left-divisible. Conversely, suppose $RA = B$ and $TB = A$. Let $UA = H$ and $VB = G$ be the first canonical forms of A and B , respectively. Then $A = U^{-1}H$ and $B = V^{-1}G$ imply $RU^{-1}H = V^{-1}G$ and $TV^{-1}G = U^{-1}H$. Whence, $PH = G$ and $QG = H$, where $P = VRU^{-1}$ and $Q = UTV^{-1}$; that is, H and G are mutually left-divisible. In proving the uniqueness of the first canonical form H , we showed that $PH = G$ and $QG = H$, where P and Q are unimodular, imply $H = G$. However, the unimodularity of P and Q was not used anywhere in this proof. Hence, we established at that point also that if H and G are mutually left-divisible, then $H = G$. This enables us to state the following result.

THEOREM 4.1. *A necessary and sufficient condition that two matrices A and B of \mathfrak{M} be mutually left-divisible is that they be left-associates.*

Let us now consider the matrix

$$\begin{bmatrix} A & 0 \\ B & 0 \end{bmatrix}$$

of order $2n$. Then there exists a unimodular matrix X of order $2n$, which we write in the form of $n \times n$ blocks, such that

$$\begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \begin{bmatrix} A & 0 \\ B & 0 \end{bmatrix} = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$$

is the first canonical form of the above matrix. Thus

$$X_{11}A + X_{12}B = D$$

so that every c.r.d. of A and B is a right divisor of D . Since X is unimodular there exists a matrix $Y = X^{-1}$ such that

$$\begin{bmatrix} A & 0 \\ B & 0 \end{bmatrix} = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix};$$

whence $A = Y_{11}D$, $B = Y_{21}D$, so that D is a c.r.d. of A and B . Hence D is a g.c.r.d. of A and B .

The matrix $M = X_{21}A = X_{22}B$ is a c.l.m. of A and B . Using an argument due to Stewart (6), we are able to show that M is the l.c.l.m. of A and B when $D = I$. To do this, let $M_1 = UA = VB$ be any other c.l.m. of A and B . We can then write the following equations:

$$\begin{bmatrix} X_{11} & X_{12} \\ U & V \end{bmatrix} \begin{bmatrix} A & 0 \\ B & 0 \end{bmatrix} = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix},$$

$$\begin{bmatrix} X_{11} & X_{12} \\ U & V \end{bmatrix} \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}.$$

Consider the most general solution of the equation

$$\begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}.$$

Here Z_{12} and Z_{22} are arbitrary, but Z_{11} and Z_{21} must be chosen so that

$$Z_{11}D = D, \quad Z_{21}D = 0.$$

Subject to these conditions the following equations must hold:

$$\begin{bmatrix} X_{11} & X_{12} \\ U & V \end{bmatrix} \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix},$$

$$\begin{bmatrix} X_{11} & X_{12} \\ U & V \end{bmatrix} = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}.$$

In particular, it appears that U has the form

$$U = Z_{21}X_{11} + Z_{22}X_{21}.$$

But if $D = I$ the only solution of $Z_{21}D = 0$ is $Z_{21} = 0$. Hence it follows in this case that $U = Z_{22}X_{21}$; then from $M_1 = UA = Z_{22}X_{21}A = Z_{22}M$ it follows that $M = X_{21}A = X_{22}B$ is indeed a l.c.l.m. of A and B . These results are stated in the following theorem.

THEOREM 4.2. *In the matrix equation*

$$\begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \begin{bmatrix} A & 0 \\ B & 0 \end{bmatrix} = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix},$$

written in the form of $n \times n$ blocks, where X is unimodular, the matrix D is in all cases a g.c.r.d. of A and B ; if $D = I$, then the matrix $M = X_{21}A = X_{22}B$ is a l.c.l.m. of A and B .

THEOREM 4.3. *The g.c.r.d. D and the l.c.l.m. M of two matrices A and B are uniquely determined up to unimodular left factors.*

If D and D_1 are two g.c.r.d.'s of A and B , then each is a c.l.m. of the other, say $D = UD_1$ and $D_1 = VD$. Then by theorem 4.1, D and D_1 are left-associates.

If M and M_1 are two l.c.l.m.'s of A and B , then each is a common right divisor of the other, say $M_1 = UM$ and $M = VM_1$. Then by Theorem 4.1, M and M_1 are left-associates.

5. Conclusion. The analogy of our results to the corresponding ones for the classical case seems remarkable when one considers that a principal ideal ring contains no proper divisors of 0, whereas every element of a Boolean ring except 1 and 0 is a proper divisor of 0. Finally we mention that the restriction to square matrices was inessential.

REFERENCES

1. E. T. Bell, *Arithmetic of logic*, Trans. Amer. Math. Soc., 29 (1927), 597-611.
2. G. Birkhoff, *Lattice Theory*, Amer. Math. Soc. Colloq. Publ., 25 (rev. ed., New York, 1948).
3. I. Kaplansky, *Elementary divisors and modules*, Trans. Amer. Math. Soc., 66 (1949), 464-491.
4. C. C. MacDuffee, *Matrices with elements in a principal ideal ring*, Bull. Amer. Math. Soc., 39 (1933), 564-584.
5. E. Steinitz, *Rechteckige Systeme und Moduln in algebraischen Zahlkörpern*, Math. Ann., 71 (1911), 328-354, and 72 (1912), 297-345.
6. B. M. Stewart, *A note on least common left multiples*, Bull. Amer. Math. Soc., 55 (1949), 587-591.

University of Florida

CHARACTERISTIC POLYNOMIALS

HANS SCHNEIDER

Introduction. Let F be a field and let V be a finite dimensional vector space over F which is also a module over the ring $F[\mathbf{a}]$. Here \mathbf{a} may lie in any extension ring of F . We do not assume, as yet, that V is a faithful module, so that \mathbf{a} need not be a linear transformation on V . It is known that by means of a decomposition of V into cyclic $F[\mathbf{a}]$ -modules we may obtain a definition of the characteristic polynomial of \mathbf{a} on V which does not involve determinants. In this note we shall give another non-determinantal definition of the characteristic polynomial. Instead of considering a single module V , we shall accordingly study the set of all finite dimensional $F[\mathbf{a}]$ -modules and mappings of this set into monic polynomials with coefficients in F . Admittedly our procedure does not yield the theory of the elementary divisors of \mathbf{a} , but it has certain advantages. First, all questions of uniqueness are settled immediately by the Jordan-Hölder theorem. Secondly, it is possible to derive some classical results, usually proved using determinants, without excessive labour. To illustrate the use of our method we shall complete and generalise some results due to Goldhaber (2) and Osborne (5).

1. A principal ideal in a semigroup of mappings. Let \mathfrak{B} be the set of all finite dimensional $F[\mathbf{a}]$ -modules, and let \mathfrak{M} be the multiplicative semigroup of (non-zero) monic polynomials with coefficients in F . Let \mathfrak{T} be the set of mappings of \mathfrak{B} into \mathfrak{M} . We shall assume throughout that if ψ is a mapping of \mathfrak{T} , and if V_1 and V_2 are isomorphic modules of \mathfrak{B} , then $\psi(V_1, t) = \psi(V_2, t)$. The set \mathfrak{T} becomes a semigroup if we define multiplication in the obvious way, viz. $\psi_1\psi_2(V, t) = \psi_1(V, t) \cdot \psi_2(V, t)$ for all V in \mathfrak{B} . The subset \mathfrak{S} of \mathfrak{T} consisting of all mappings satisfying

$$(1) \quad \psi(V, t) = \psi(V/Z, t) \cdot \psi(Z, t)$$

for all $V \in \mathfrak{B}$ and all submodules Z of V , is also a semigroup. Now let μ be any mapping of \mathfrak{T} such that $\mu((0), t) = 1$ and

$$(2) \quad \mu(V, t) \text{ divides } \mu(V/Z, t) \cdot \mu(Z, t)$$

for all $V \in \mathfrak{B}$ and all submodules Z of V . Let \mathfrak{G} be the ideal in the semigroup \mathfrak{S} consisting of all ψ of \mathfrak{S} divisible by μ in \mathfrak{T} i.e. let $\mathfrak{G} = \mu\mathfrak{T} \cap \mathfrak{S}$.

Then an element ψ of \mathfrak{T} belongs to \mathfrak{G} if and only if ψ satisfies (1) and

$$(3) \quad \mu(V, t) \text{ divides } \psi(V, t)$$

for all $V \in \mathfrak{B}$.

Received March 1, 1956. I should like to express my gratitude to Dr. I. T. Adamson and Dr. M. P. Drazin for discussing with me the results of this paper.

Let $V = V_0 \supset \dots \supset V_s = (0)$ be a composition series for the $F[\mathfrak{a}]$ -module V . Then it is an immediate consequence of the Jordan-Hölder theorem that the polynomial

$$(4) \quad \chi(V, t) = \prod_{i=1}^s \mu(V_{i-1}/V_i, t) \cdot \mu((0), t)$$

depends only on the module V and not on the composition series we use in the definition. (The last factor $\mu((0), t) = 1$ has been put in to cover the case $V = (0)$, yielding $\chi((0), t) = \mu((0), t) = 1$. In some of the arguments below it has been assumed that $V \neq (0)$, the case $V = (0)$ being trivial.)

This remark allows us to think of χ as a mapping of \mathfrak{B} into \mathfrak{M} , i.e. as an element of the semigroup \mathfrak{T} . The semigroup \mathfrak{S} is not a principal ideal semigroup. But we shall prove

LEMMA 1. *The mapping χ defined by (4) is the unique generator of the ideal \mathfrak{C} in \mathfrak{S} .*

Proof. We shall first show that $\chi \in \mathfrak{S}$. If Z is any submodule of V then there is a composition series $V = V'_0 \supset \dots \supset V'_s = (0)$ in which Z occurs, say $Z = V'_r$. Thus it follows from our remarks about the Jordan-Hölder theorem, and since $(V_{i-1}'/Z)/(V_i'/Z)$ is isomorphic to V_{i-1}'/V_i' that

$$\chi(V/Z, t) \chi(Z, t) = \prod_{i=1}^r \mu(V'_{i-1}/V'_i, t) \prod_{i=r+1}^s \mu(V'_{i-1}/V'_i, t) = \chi(V, t)$$

whence $\chi \in \mathfrak{S}$.

It follows from (3) applied to the factor modules that

$$\prod_{i=1}^s \mu(V_{i-1}/V_i, t)$$

divides $\chi(V, t)$. Hence by (2), $\mu(V, t)$ divides $\chi(V, t)$. Thus μ divides χ , and we deduce that $(\chi) \subseteq \mathfrak{C}$.

To prove the reverse inclusion let us suppose that $\psi \in \mathfrak{C}$. Then for any $V \in \mathfrak{B}$, we have

$$\psi(V, t) = \prod_{i=1}^s \psi(V_{i-1}/V_i, t)$$

by (1). But, by (3), $\mu(V_{i-1}/V_i, t)$ divides $\psi(V_{i-1}/V_i, t)$, whence $\chi(V, t)$ divides $\psi(V, t)$. It follows that χ divides ψ in \mathfrak{T} , say $\psi = \chi\psi^*$. To show that ψ^* belongs to \mathfrak{S} we observe that for all submodules Z of V

$$\begin{aligned} \chi(V, t) \psi^*(V, t) &= \psi(V, t) = \psi(V/Z, t) \psi(Z, t) \\ &= \chi(V/Z, t) \chi(Z, t) \psi^*(V/Z, t) \psi^*(Z, t) = \chi(V, t) \psi^*(V/Z, t) \psi^*(Z, t) \end{aligned}$$

whence ψ^* satisfies (1) since $\chi(V, t)$ is non-zero. We have now proved that $\psi \in (\chi)$, and this implies that $\mathfrak{C} \subseteq (\chi)$. It follows that $(\chi) = \mathfrak{C}$, and we conclude that χ is a generator of \mathfrak{C} .

To prove uniqueness, let us now suppose that $\mathfrak{E} = (\chi) = (\chi^*)$. Then $\chi^* = \chi\psi$ and also $\chi = \chi^*\psi^*$. We obtain that $\chi = \chi\psi\psi^*$, and it follows that $\psi\psi^* = \eta$, where $\eta(V, t) = 1$, for all $V \in \mathfrak{B}$. The mapping η is the identity of \mathfrak{E} (and also, of course, of \mathfrak{T}), and since η is the only inverse in \mathfrak{T} , it follows that $\psi = \eta$, whence $\chi = \chi^*$, and the lemma is proved.

COROLLARY. *If ψ belongs to the ideal \mathfrak{E} , and $\psi(V, t)$ has the same degree as $\chi(V, t)$ for all V in \mathfrak{B} , then $\chi = \psi$.*

Proof. Clearly $\psi = \chi\psi^*$, where, for all V in \mathfrak{B} , $\psi^*(V, t)$ is a monic polynomial of degree 0, and the only such polynomial is 1.

2. The definition of the characteristic polynomial. As is well known, the polynomials $p(t)$ in $F[t]$ for which $p(\mathbf{a})V = 0$ form an ideal in $F[t]$. By a slight extension of the usual terminology, we shall call the monic generator of this ideal the *minimum polynomial of \mathbf{a} on V* , and denote it by $\mu_{\mathbf{a}}(V, t)$. Thus $\mu_{\mathbf{a}}$ is a mapping of \mathfrak{B} into \mathfrak{M} such that $\mu_{\mathbf{a}}((0), t) = 1$ and which satisfies (2) since

$$\mu_{\mathbf{a}}(Z, \mathbf{a}) \cdot \mu_{\mathbf{a}}(V/Z, \mathbf{a}) \quad V \subseteq \mu_{\mathbf{a}}(Z, \mathbf{a}) \quad Z = (0),$$

when Z is any submodule of V . It follows, therefore, that the ideal $\mathfrak{E}_{\mathbf{a}}$, which consists of all $\psi \in \mathfrak{T}$ divisible by $\mu_{\mathbf{a}}$, and which we shall call the *characteristic ideal of \mathbf{a} in \mathfrak{E}* , is a principal ideal with a unique generator. Thus we may make the following definition:

Definition. Let $V \in \mathfrak{B}$. Then the *characteristic polynomial of \mathbf{a} on V* is $\chi_{\mathbf{a}}(V, t)$, where $\chi_{\mathbf{a}}$ is the unique generator of the characteristic ideal $\mathfrak{E}_{\mathbf{a}} = \mu_{\mathbf{a}}\mathfrak{T} \cap \mathfrak{E}$ of \mathbf{a} .

By virtue of Lemma 1, we obtain immediately

THEOREM 1. *Let $V \in \mathfrak{B}$, and let $V = V_0 \supset \dots \supset V_s = (0)$ be a composition series for V . Then the characteristic polynomial $\chi_{\mathbf{a}}(V, t)$ of \mathbf{a} on V is*

$$(5) \quad \chi_{\mathbf{a}}(V, t) = \prod_{i=1}^s \mu_{\mathbf{a}}(V_{i-1}/V_i, t) \cdot \mu_{\mathbf{a}}((0), t).$$

COROLLARY 1. *For each V in \mathfrak{B} , the degree of $\chi_{\mathbf{a}}(V, t)$ equals the dimension of V .*

Proof. In view of (5), we need only prove that the degree of $\mu_{\mathbf{a}}(Z, t)$ equals the dimension of Z , when Z is an irreducible $F[\mathbf{a}]$ -module. This is well-known if Z is faithful over $F[\mathbf{a}]$. If Z is not faithful over $F[\mathbf{a}]$, then we must have $\mathbf{a}Z = 0$. The dimension of Z must equal 1, and $\mu_{\mathbf{a}}(Z, t) = t$.

COROLLARY 2. *If ψ belongs to the characteristic ideal of \mathbf{a} , and if the degree of $\psi(V, t)$ equals the dimension of V for each $V \in \mathfrak{B}$, then $\psi = \chi_{\mathbf{a}}$.*

This corollary follows from the previous one and from the corollary to Lemma 1.

We shall now show that our definition of the characteristic polynomial leads to the usual one in the case of matrices. For any basis for V we obtain a matrix of \mathbf{a} on V in the usual way. Since we have not assumed that \mathbf{a} is a linear transformation on V , a non-zero \mathbf{a} may have a zero matrix on V . The determinant of a matrix A will be denoted by $|A|$.

THEOREM 2. *If A is any matrix of \mathbf{a} on V and I is the unit matrix, then*

$$(6) \quad \chi_{\mathbf{a}}(V, t) = |tI - A|.$$

Proof. Any two basis for V yield similar matrices of \mathbf{a} on V . Hence the determinant $|tI - A|$ depends on V only and not on the basis used to obtain A . Thus we may define a mapping of \mathfrak{T} by setting $\psi(V, t) = |tI - A|$.

Let Z be any submodule of V . We choose a basis for Z and complete it to a basis for V . With respect to this basis the matrix of \mathbf{a} on V is

$$B = \begin{bmatrix} B_1 & B_{12} \\ 0 & B_2 \end{bmatrix},$$

where B_1 is a matrix of \mathbf{a} on Z , and B_2 is a matrix of \mathbf{a} on V/Z .

Hence

$$\psi(V, t) = |tI - B| = |tI - B_1| |tI - B_2| = \psi(Z, t) \psi(V/Z, t);$$

and so ψ belongs to \mathfrak{S} .

It is easily verified that $p(A) = 0$ is equivalent to $p(\mathbf{a})V = 0$, whence the minimum polynomial of the matrix A is just $\mu_{\mathbf{a}}(V, t)$. Thus by the classical Cayley-Hamilton theorem applied to the determinant $|tI - A|$, the polynomial $\mu_{\mathbf{a}}(V, t)$ divides $\psi(V, t)$. It follows that ψ belongs to the characteristic ideal of \mathbf{a} . The degree of $\psi(V, t)$ is clearly equal to the dimension of V . We may now use Corollary 2 to Theorem 1 to conclude that $\chi_{\mathbf{a}}(V, t) = \psi(V, t) = |tI - A|$.

We may remark that it is possible to obtain an extension of some of the above results to the case of $F[\mathbf{a}]$ -modules of infinite dimension over F , in which case V has no composition series and $\psi(V, t) = 0$.

3. Characteristic polynomials with a common factor. In this section we shall again assume that the $F[\mathbf{a}]$ -module V is finite dimensional. If Z is a submodule of V , then $\mu_{\mathbf{a}}(V, t)$ is clearly divisible by both $\mu_{\mathbf{a}}(V/Z, t)$ and $\mu_{\mathbf{a}}(Z, t)$. But the product $\mu_{\mathbf{a}}(V/Z, t) \cdot \mu_{\mathbf{a}}(Z, t)$ is divisible by $\mu_{\mathbf{a}}(V, t)$. It follows by (5) that $\chi_{\mathbf{a}}(V, t)$ divides a power of $\mu_{\mathbf{a}}(V, t)$. Thus we have proved from our definitions the very well-known result that every irreducible factor of $\chi_{\mathbf{a}}(V, t)$ is also a factor of $\mu_{\mathbf{a}}(V, t)$. This enables us to prove the next lemma.

LEMMA 2. *There exists a composition series for V with the factor modules appearing in any order.*

Proof. Let $p(t)$ be any irreducible factor of $\chi_{\mathbf{a}}(V, t)$ and therefore also of $\mu_{\mathbf{a}}(V, t)$. Let $q(t) = \mu_{\mathbf{a}}(V, t)/p(t)$, and let Z be the submodule of V consisting of all $v \in V$ for which $q(\mathbf{a})v = 0$. Then $p(\mathbf{a})V \subseteq Z$, and so $\mu_{\mathbf{a}}(V/Z, t)$ divides

$p(t)$. But Z is a proper submodule of V , whence $\mu_a(V/Z, t) \neq 1$. We deduce that the minimum polynomial of \mathbf{a} on V/Z and any of its irreducible submodules is $p(t)$. Since we may start a composition series for V with a composition series for V/Z , the lemma follows.

LEMMA 3. *Let Z be a submodule of the $F[\mathbf{a}]$ -module V , and let Y be a submodule of the $F[\mathbf{a}]$ -module W . Let λ_1 be a vector space isomorphism of Y onto Z . If V/Z and W/Y are irreducible submodules for which $\mu_a(V/Z, t) = \mu_a(W/Y, t)$, then there exists an extension of λ_1 to a vector space isomorphism λ of W onto V such that $\mathbf{a}w^\lambda - (\mathbf{b}w)^\lambda \in Z$, for all $w \in W$.*

Proof. The conditions on V/Z and W/Y imply that there is a vector space isomorphism κ of W/Y onto V/Z such that $\mathbf{a}(w + Y)^\kappa = (\mathbf{b}(w + Y))^\kappa$ for all $w \in W$. Let Z' be a subspace of V complementary to Z , and let Y' be a subspace of W complementary to Y . We shall now define a mapping of W into V as follows.

We let λ coincide with λ_1 on Y . If $w \in Y'$ we let $v = w^\lambda$ be the unique element of Z' for which $v + Z = (w + Y)^\kappa$. We then extend λ linearly from $Y \cup Y'$ to $W = Y + Y'$. Then λ is an isomorphism onto V , and the Lemma follows since, for all $w \in Y'$, we have

$$\begin{aligned} \mathbf{a}w^\lambda + Z &= \mathbf{a}(w^\lambda + Z) = \mathbf{a}(w + Y)^\kappa = (\mathbf{b}(w + Y))^\kappa = (\mathbf{b}w + Y)^\kappa \\ &= (\mathbf{b}w)^\lambda + Z. \end{aligned}$$

Let $\mathfrak{L}(V)$ be the algebra over F of linear transformations on the finite dimensional vector space V . If W is also a vector space over F , and λ is a vector space isomorphism of W onto V , then λ induces an isomorphism ρ of $\mathfrak{L}(W)$ onto $\mathfrak{L}(V)$, which may be defined by $\mathbf{b}^*w^\lambda = (\mathbf{b}w)^\lambda$, for all $w \in W$. Conversely, let ρ be an isomorphism of $\mathfrak{L}(W)$ onto $\mathfrak{L}(V)$. If we consider F as a subalgebra of both $\mathfrak{L}(W)$ and $\mathfrak{L}(V)$ in the normal way, then elements of F are left fixed by ρ . Hence it can be shown that ρ is induced by an isomorphism λ of W onto V (4, p. 237). It is this result which almost immediately yields the following lemma.

LEMMA 4. *Let ρ be an isomorphism of the algebra of linear transformations $\mathfrak{L}(W)$ onto $\mathfrak{L}(V)$. If $\mathbf{b} \in \mathfrak{L}(W)$ and $\mathbf{c} = \mathbf{b}^\rho$, then $\chi_b(W, t) = \chi_c(V, t)$.*

Proof. Let λ be the isomorphism of the vector space W onto V which induces ρ . Then any composition series for the $F[\mathbf{b}]$ -module W is mapped by λ into a composition series for the $F[\mathbf{c}]$ -module V . It is easy to see that the minimum polynomial of \mathbf{b} on a factor module of the first series equals the minimum polynomial of \mathbf{c} on the corresponding factor module of the second series. The lemma now follows from Theorem 1.

This lemma is rather less trivial than may appear at first sight. For, if ρ is an isomorphism merely of the subalgebra $F[\mathbf{b}]$ of $\mathfrak{L}(W)$ into $\mathfrak{L}(V)$, and $\mathbf{c} = \mathbf{b}^\rho$, then we can only conclude that $\mu_b(W, t) = \mu_c(V, t)$.

Goldhaber (2) and Goldhaber and Whaples (3) have proved that if A and B are square matrices with coefficients in F such that there exists a non-singular matrix P for which $N = A - PBP^{-1}$ lies in the radical of $F[A, PBP^{-1}]$ then $|tI - A| = |tI - B|$. This result is generally known as Goldhaber's lemma. By means of the theory of canonical matrices and under the assumption that F is an infinite perfect field Osborne (5) has proved this lemma together with its converse. From now on we shall assume that \mathbf{a} and \mathbf{b} are linear transformations on the finite dimensional vector spaces V and W respectively, and we shall prove a theorem equivalent to Osborne's without any restriction on the field F .

THEOREM 3. *Let \mathbf{a} and \mathbf{b} be linear transformations on the finite dimensional vector spaces V and W respectively. Then the characteristic polynomial of \mathbf{a} on V equals the characteristic polynomial of \mathbf{b} on W if and only if there is an isomorphism ρ of $\mathfrak{L}(W)$ onto $\mathfrak{L}(V)$ such that $\mathbf{n} = \mathbf{a} - \mathbf{b}^\rho$ lies in the radical of $F[\mathbf{a}, \mathbf{b}^\rho]$.*

Proof. If ρ is an isomorphism of $\mathfrak{L}(W)$ onto $\mathfrak{L}(V)$, and $\mathbf{c} = \mathbf{b}^\rho$, then by Lemma 4 we need only prove that $\chi_{\mathbf{a}}(V, t) = \chi_{\mathbf{c}}(V, t)$ when $\mathbf{n} = \mathbf{a} - \mathbf{c}$ lies in the radical of

$$F[\mathbf{a}, \mathbf{c}] = F[\mathbf{a}, \mathbf{n}] = F[\mathbf{c}, \mathbf{n}].$$

Let Z be an irreducible $F[\mathbf{a}, \mathbf{c}]$ -module. Then $\mathbf{n}Z = (0)$, whence Z is irreducible also over $F[\mathbf{a}]$ and $F[\mathbf{c}]$. Further, $\mu_{\mathbf{a}}(Z, t) = \mu_{\mathbf{c}}(Z, t)$, since for a polynomial $p(t)$ in $F[t]$ the equality $p(\mathbf{c})Z = 0$ implies $(p(\mathbf{a}) + \mathbf{n}')Z = 0$, where \mathbf{n}' belongs to the radical of $F[\mathbf{a}, \mathbf{c}]$, whence $p(\mathbf{a})Z = 0$; and conversely. Let

$$V = V_0 \supset \dots \supset V_s = (0)$$

be a composition series for the $F[\mathbf{a}, \mathbf{c}]$ -module V . It follows from the remarks we have just made that this series is also a composition series for V as an $F[\mathbf{a}]$ and $F[\mathbf{c}]$ -module, and that

$$\mu_{\mathbf{a}}(V_{i-1}/V_i, t) = \mu_{\mathbf{c}}(V_{i-1}/V_i, t), \quad i = 1, \dots, s.$$

Hence by Theorem 1, $\chi_{\mathbf{a}}(V, t) = \chi_{\mathbf{c}}(V, t)$.

Now let us suppose that $\chi_{\mathbf{a}}(V, t) = \chi_{\mathbf{b}}(W, t)$, and let $V = V_0 \supset \dots \supset V_s = (0)$ be a composition series for the $F[\mathbf{a}]$ -module V . We deduce from Lemma 2 that there is a composition series

$$W = W_0 \supset \dots \supset W_s = (0)$$

for the $F[\mathbf{b}]$ -module W such that $\mu_{\mathbf{b}}(W_{i-1}/W_i, t) = \mu_{\mathbf{a}}(V_{i-1}/V_i, t)$, for $i = 1, \dots, s$. Then using Lemma 3 we may prove by induction that there is a vector space isomorphism λ of W onto V which takes W_i onto V_i such that $\mathbf{a}w^\lambda - (\mathbf{b}w)^\lambda \in V_i$, whenever $w \in W_{i-1}$ ($i = 1, \dots, s$). Let ρ be the isomorphism of $\mathfrak{L}(W)$ onto $\mathfrak{L}(V)$ associated with λ . If $v = w^\lambda \in V_{i-1}$, and $\mathbf{n} = \mathbf{a} - \mathbf{b}^\rho$, then

$$\mathbf{n}v = \mathbf{a}w^\lambda - \mathbf{b}^\rho w^\lambda = \mathbf{a}w^\lambda - (\mathbf{b}w)^\lambda \in V_i.$$

Thus for any polynomial $p(t)$ in $F[t]$ we have $(p(\mathbf{a})\mathbf{n})^*V = (0)$. Since V is faithful over $F[\mathbf{a}, \mathbf{b}^*]$, we conclude that \mathbf{n} lies in the radical of $F[\mathbf{a}, \mathbf{b}^*]$.

Lemma 4 allows us to state Theorem 3 symmetrically: *The characteristic polynomials $\chi_a(V, t)$ and $\chi_b(W, t)$ are equal if and only if there exist isomorphisms σ and τ of $\mathfrak{L}(V)$ and $\mathfrak{L}(W)$ respectively onto an algebra of linear transformations $\mathfrak{L}(X)$ such that $\mathbf{n} = \mathbf{a}^* - \mathbf{b}^*$ lies in the radical of $F[\mathbf{a}^*, \mathbf{b}^*]$.*

Let $\mathfrak{Z}(Z, V)$ be the subalgebra of $\mathfrak{L}(V)$ consisting of all linear transformations $\mathbf{d} \in \mathfrak{L}(V)$ for which $\mathbf{d}Z \subseteq Z$. We note that $\mathbf{d} \in \mathfrak{Z}(Z, V)$ if and only if Z is an $F[\mathbf{d}]$ -module. The natural homomorphism $\mathbf{d} \rightarrow \mathbf{d}'$ of $\mathfrak{Z}(Z, V)$ onto $\mathfrak{L}(Z)$ is defined by $\mathbf{d}'v = \mathbf{d}v$, for all $v \in Z$. Its kernel K_1 consists of all $\mathbf{d} \in \mathfrak{Z}(Z, V)$ such that $\mathbf{d}Z = (0)$. Clearly $\chi_{a'}(Z, t) = \chi_a(Z, t)$. There is also a natural homomorphism of $\mathfrak{Z}(Z, V)$ onto $\mathfrak{L}(V/Z)$, defined by $\mathbf{d} \rightarrow \mathbf{d}''$ where $\mathbf{d}''(v + Z) = \mathbf{d}v + Z$ for all $v \in V$. Its kernel K_2 consists of all $\mathbf{d} \in \mathfrak{Z}(Z, V)$ such that $\mathbf{d}V \subseteq Z$. Again $\chi_{a''}(V/Z, t) = \chi_a(V/Z, t)$.

LEMMA 5. *Let X be a vector space of dimension r over F , and let $p(t)$ be a monic polynomial of degree r in $F[t]$. Then $p(t)$ divides $\chi_a(V, t)$ if and only if there exists an $F[\mathbf{a}]$ -submodule Z contained in V and a homomorphism σ of $\mathfrak{Z}(Z, V)$ onto $\mathfrak{L}(X)$ such that $\chi_c(X, t) = p(t)$, where $\mathbf{c} = \mathbf{a}^*$.*

Proof. Let $p(t)$ be a factor of $\chi_a(V, t)$. By Lemma 2 there exists an $F[\mathbf{a}]$ -module Z contained in V for which $\chi_a(Z, t) = p(t)$, and by Corollary 1 to Theorem 1 the dimension on Z is r . Thus we may define the homomorphism σ of $\mathfrak{Z}(Z, V)$ onto $\mathfrak{L}(X)$ to be the composed map of the natural homomorphism $\mathbf{d} \rightarrow \mathbf{d}'$ of $\mathfrak{Z}(Z, V)$ onto $\mathfrak{L}(Z)$, and any isomorphism of $\mathfrak{L}(Z)$ onto $\mathfrak{L}(X)$. Then putting $\mathbf{c} = \mathbf{a}^*$, we have

$$\chi_c(X, t) = \chi_{a'}(Z, t) = \chi_a(Z, t) = p(t),$$

by virtue of Lemma 4.

Conversely, let σ be a homomorphism of $\mathfrak{Z}(Z, V)$ onto $\mathfrak{L}(X)$ for which $\chi_c(X, t) = p(t)$. Using the simplicity of $\mathfrak{L}(X)$ it may be shown that the kernel K of σ is $\mathfrak{Z}(Z, V)$, K_1 or K_2 . If $K = \mathfrak{Z}(Z, V)$, then $\mathfrak{L}(X) = (0)$, whence $X = (0)$ and $p(t) = 1$. If $K = K_1$, then $\mathfrak{L}(Z)$ is isomorphic to $\mathfrak{L}(X)$ under an isomorphism which takes \mathbf{a}' onto $\mathbf{c} = \mathbf{a}^*$. In this case we deduce that $p(t) = \chi_{a'}(Z, t) = \chi_a(Z, t)$, and $\chi_a(Z, t)$ divides $\chi_a(V, t)$. If $K = K_2$, then $\mathfrak{L}(V/Z)$ is isomorphic to $\mathfrak{L}(X)$ under an isomorphism taking \mathbf{a}'' onto \mathbf{c} , and so

$$p(t) = \chi_{a''}(V/Z, t) = \chi_a(V/Z, t),$$

which again divides $\chi_a(V, t)$.

By combining the symmetric form of Theorem 3 and Lemma 5 we immediately obtain a generalisation of Theorem 3.

THEOREM 4. Let \mathbf{a} and \mathbf{b} be linear transformations on the vector spaces V and W respectively. Let X be a vector space of dimension r over F . Then the characteristic polynomials $\chi_{\mathbf{a}}(V, t)$ and $\chi_{\mathbf{b}}(W, t)$ have a common factor of degree r if and only if there exist an $F[\mathbf{a}]$ -module Z contained in V and a homomorphism σ of $\mathfrak{Z}(Z, V)$ onto $\mathfrak{Z}(X)$, an $F[\mathbf{b}]$ -module Y contained in W and a homomorphism τ of $\mathfrak{Z}(Y, W)$ onto $\mathfrak{Z}(X)$, such that $\mathbf{n} = \mathbf{a}^\sigma - \mathbf{b}^\tau$ lies in the radical of $F[\mathbf{a}^\sigma, \mathbf{b}^\tau]$.

Finally, we claim that some of the results of Goddard and Schneider (1) and other results on characteristic polynomials may be derived from Theorem 4.

REFERENCES

1. L. S. Goddard and H. Schneider, *Matrices with a non-zero commutator*, Proc. Cambridge Phil. Soc. 51 (1955), 551-553.
2. J. K. Goldhaber, *The homomorphic mapping of certain matrix algebras onto rings of diagonal matrices*, Can. J. Math. 4 (1952), 31-42.
3. J. K. Goldhaber and G. Whaples, *On some matrix theorems of Frobenius and McCoy*, Can. J. Math. 5 (1953), 332-335.
4. N. Jacobson, *Lectures on abstract algebra*, vol. 2 (New York, 1953).
5. E. E. Osborne, *On matrices having the same characteristic equation*, Pacific J. Math. 2 (1952), 227-230.

Queen's University, Belfast and
Washington State College
Pullman, Wash.

SOME THEOREMS ABOUT $p_r(n)$

MORRIS NEWMAN

Introduction. If n is a non-negative integer, define $p_r(n)$ as the coefficient of x^n in

$$\prod_{n=1}^{\infty} (1 - x^n)^r;$$

otherwise define $p_r(n)$ as 0. In a recent paper (1) the author has proved that if r has any of the values 2, 4, 6, 8, 10, 14, 26 and p is a prime > 3 such that $r(p+1) \equiv 0 \pmod{24}$, then

$$(1) \quad p_r(np + \Delta) = (-p)^{\frac{1}{2}r-1} p_r\left(\frac{n}{p}\right), \quad \Delta = r(p^2 - 1)/24,$$

where n is an arbitrary integer.

In this note we wish to point out one or two additional facts implied by identity (1). The first remark is that (1) furnishes a simple, uniform proof of the Ramanujan congruences for partitions modulo 5, 7, 11, and a general congruence will be proved. The second is that for the values of r indicated, $p_r(n)$ is zero for arbitrarily long strings of consecutive values of n . Finally, some additional theorems not covered by (1) will be given without proof.

In what follows all products will be extended from 1 to ∞ and all sums from 0 to ∞ , unless otherwise indicated.

THEOREM 1. Let $r = 4, 6, 8, 10, 14, 26$. Let p be a prime greater than 3 such that $r(p+1) \equiv 0 \pmod{24}$, and set $\Delta = r(p^2 - 1)/24$. Then if $R \equiv r \pmod{p}$ and $n \equiv \Delta \pmod{p}$,

$$(2) \quad p_R(n) \equiv 0 \pmod{p}.$$

Proof. Set $R = Qp + r$. Then

$$\begin{aligned} \sum p_r(n)x^n &= \prod (1 - x^n)^R = \prod (1 - x^n)^{Qp+r} \\ &\equiv \prod (1 - x^{np})^Q (1 - x^n)^r \pmod{p}. \end{aligned}$$

Thus

$$\sum p_R(n)x^n \equiv \sum p_Q\left(\frac{n}{p}\right)x^n \sum p_r(n)x^n \pmod{p},$$

and so

$$p_R(n) \equiv \sum_{j=0}^n p_Q\left(\frac{j}{p}\right) p_r(n-j) \pmod{p},$$

Received April 3, 1956. The preparation of this paper was supported (in part) by the Office of Naval Research.

or

$$p_R(n) \equiv \sum_{0 < j < \frac{n}{p}} p_Q(j) p_r(n - pj) \pmod{p}.$$

Now (1) implies that for $r > 2$ and $n \equiv \Delta \pmod{p}$, $p_r(n - pj) \equiv 0 \pmod{p}$. Thus $p_R(n) \equiv 0 \pmod{p}$, and so (2) is proved.

If we now note that for $R = -1$ the choices $r = 4$, $p = 5$; $r = 6$, $p = 7$; and $r = 10$, $p = 11$ (all with $Q = -1$) are permissible, and that for these values $\Delta = 4, 12, 50$ respectively, then the Ramanujan congruences $p(5n + 4) \equiv 0 \pmod{5}$, $p(7n + 5) \equiv 0 \pmod{7}$, $p(11n + 6) \equiv 0 \pmod{11}$ follow as a corollary, since $12 \equiv 5 \pmod{7}$ and $50 \equiv 6 \pmod{11}$.

We go on now to the second remark. We first prove the following lemma.

LEMMA 1. Let a_1, a_2, \dots, a_{n+1} be non-zero pairwise relatively prime integers, and let c_1, c_2, \dots, c_n be arbitrary integers. Then the simultaneous diophantine equations

$$\begin{aligned} (3) \quad & a_1 x_1 - a_2 x_2 = c_1, \\ & a_2 x_2 - a_3 x_3 = c_2, \\ & \dots \\ & a_n x_n - a_{n+1} x_{n+1} = c_n, \end{aligned}$$

always have infinitely many solutions.

Proof. Put $T = a_{n+1} x_{n+1}$, $C_i = c_i + c_{i+1} + \dots + c_n$, $1 \leq i \leq n$. Then by summing the rows of (3) 1, 2, ... at a time beginning with the last, we find that the system (3) is equivalent to the system

$$\begin{aligned} a_i x_i &= c_i + T, & 1 \leq i \leq n, \\ a_{n+1} x_{n+1} &= T. \end{aligned}$$

Since the a 's are pairwise relatively prime, the Chinese remainder theorem assures us of the existence of an integer C such that

$$C \equiv -C_i \pmod{a_i}, \quad 1 \leq i \leq n,$$

and

$$C \equiv 0 \pmod{a_{n+1}}.$$

Put $T = C + Ax$, where $A = a_1 a_2 \dots a_{n+1}$. Then (3) has the solution

$$\begin{aligned} x_i &= \frac{C + C_i}{a_i} + \frac{Ax}{a_i}, & 1 \leq i \leq n, \\ x_{n+1} &= \frac{C}{a_{n+1}} + \frac{A}{a_{n+1}} x, \end{aligned}$$

where x is arbitrary. Thus Lemma 1 is proved.

If we now notice, for example, that $p_r(np^2 + p + \Delta) = 0$ (obtained by replacing n by $np + 1$ in (1)) and that any two distinct primes p are relatively prime, we see that Lemma 1 implies

THEOREM 2. For $r = 2, 4, 6, 8, 10, 14, 26$ $p_r(n)$ vanishes for arbitrarily long strings of consecutive values of n , arbitrarily many in number.

We remark that the same is true for $p_1(n)$, $p_3(n)$, because of the classical identities

$$\prod (1 - x^n) = \sum_{n=-\infty}^{\infty} (-1)^n x^{\frac{1}{2}(3n^2+n)}$$

$$\prod (1 - x^n)^3 = \sum (-1)^n (2n+1) x^{\frac{1}{2}(n^2+n)},$$

due respectively to Euler and Jacobi.

Finally, we state without proof some additional identities derivable in the same way that (1) was derived in (1); p is a prime in what follows.

$$(4) \quad p_2\left(np + \frac{1}{12}(p^2 - 1)\right) = (-1)^{\frac{1}{2}(p+1)} p_2\left(\frac{n}{p}\right), \quad p \not\equiv 1 \pmod{12}, \quad p > 3.$$

$$(5) \quad p_6(3n + 2) = 9p_6\left(\frac{1}{3}n\right).$$

$$(6) \quad p_8(2n + 1) = -8p_8\left(\frac{1}{2}n\right) \quad \text{(due to van der Pol (2))}$$

$$(7) \quad p_{10}\left(np + \frac{5}{12}(p^2 - 1)\right) = p^4 p_{10}\left(\frac{n}{p}\right), \quad p \equiv 7 \pmod{12}.$$

$$(8) \quad p_{14}\left(np + \frac{7}{12}(p^2 - 1)\right) = -p^6 p_{14}\left(\frac{n}{p}\right), \quad p \equiv 5 \pmod{12}.$$

REFERENCES

1. M. Newman, *An identity for the coefficients of certain modular forms*. J. Lond. Math. Soc., 30 (1955), 488-493.
2. B. van der Pol, *The representation of numbers as sums of eight, sixteen, and twenty-four squares*, Proc. Kon. Nederl. Akad. Wetensch. Ser. A 57 = Indagationes Math. 16 (1954), 349-361.

National Bureau of Standards,
Washington, D.C.

CLASSES OF POSITIVE DEFINITE UNIMODULAR CIRCULANTS

MORRIS NEWMAN AND OLGA TAUSKY

All matrices considered here have rational integral elements. In particular some circulants of this nature are investigated. An $n \times n$ circulant is of the form

$$C = \begin{bmatrix} c_0 & c_1 & \dots & c_{n-1} \\ c_{n-1} & c_0 & \dots & c_{n-2} \\ & & \dots & \\ c_1 & c_2 & \dots & c_0 \end{bmatrix}$$

The following result concerning positive definite unimodular circulants was obtained recently (3; 4):

Let C be a unimodular $n \times n$ circulant and assume that $C = AA'$, where A is an $n \times n$ matrix and A' its transpose. Then it follows that $C = C_1 C_1'$ where C_1 is again a circulant.

For general unimodular matrices the assumption $C = AA'$ is stronger than symmetry and positive definiteness if and only if $n \geq 8$, as was shown by Minkowski (1). The question therefore arises whether symmetry and positive definiteness suffice even for $n \geq 8$ in the theorem above; or in other words, whether a unimodular symmetric positive definite circulant is necessarily of the form AA' . (In this connection it was shown by I. Schoenberg (in a written communication) that a hermitian positive definite circulant with arbitrary complex elements is always of the form AA' where A is again a circulant).

It will be shown that the circulant M whose first row is

$$(2, 1, 0, -1, -1, -1, 0, 1)$$

is positive definite, unimodular, but not of the form AA' .

Mordell (2) showed that every symmetric positive definite unimodular 8×8 matrix which is not of the form AA' is congruent to the matrix K which corresponds to the quadratic form

$$\sum_{i=1}^8 x_i^2 + \left(\sum_{i=1}^8 x_i \right)^2 - 2x_1x_2 - 2x_5x_8.$$

The circulant M therefore is congruent to K .

Received April 12, 1956. The preparation of this paper was sponsored (in part) by the Office of Naval Research.

THEOREM 1. *The circulant M is not of the form AA' .*

Proof. Any matrix of the form AA' corresponds to a quadratic form which represents all integers if $n > 4$, but certainly represents both odd and even integers for any n . The quadratic form corresponding to M , however, represents only even integers. This proves the theorem.

That M is positive definite can be verified directly. It is no more difficult to characterize all positive definite symmetric unimodular 8×8 circulants. This is done in the following lemma.

LEMMA 1. *Any circulant C whose first row is (a_0, a_1, \dots, a_7) is unimodular, symmetric, and positive definite if and only if*

$$\begin{aligned} a_0 &= \frac{1}{2}(1+x), & a_1 &= a_7 = \frac{1}{2}y, & a_2 &= a_6 = 0, \\ a_3 &= a_5 = -\frac{1}{2}y, & a_4 &= \frac{1}{2}(1-x), \end{aligned}$$

where $x > 0$ and $x^2 - 2y^2 = 1$. (The circulant M arises from $x = 3$, $y = 2$.)

Proof. Any circulant C with first row (a_0, a_1, \dots, a_7) has the eight characteristic roots

$$\alpha_i = \sum_{j=0}^7 a_j \zeta^{ij}$$

where ζ runs through the eight roots of $x^8 - 1 = 0$. The circulant C is unimodular and positive definite if the algebraic integers α_i are real positive units. From this it follows that C is unimodular, symmetric, and positive definite if and only if

$$\begin{aligned} (1) \quad & a_0 + 2a_1 + 2a_2 + 2a_3 + a_4 = 1 & (\zeta = 1), \\ (2) \quad & a_0 - 2a_1 + 2a_2 - 2a_3 + a_4 = 1 & (\zeta = -1), \\ (3) \quad & a_0 - 2a_2 + a_4 = 1 & (\zeta^2 = -1), \\ (4) \quad & a_0 - a_4 + (a_1 - a_3)(\zeta - \zeta^3) = \epsilon_1 & (\zeta^4 = -1), \\ (5) \quad & a_0 - a_4 - (a_1 - a_3)(\zeta - \zeta^3) = \epsilon_2 & (\zeta^4 = -1), \end{aligned}$$

where ϵ_1, ϵ_2 are real and positive units.

The equations (1), (2), (3) imply that $a_2 = 0$, $a_0 + a_4 = 1$, $a_1 + a_3 = 0$. Introducing these relations and $\zeta - \zeta^3 = \pm \sqrt{2}$ for $\zeta^4 = -1$ into (4) and (5) we obtain

$$\begin{aligned} 2a_0 - 1 + 2a_1\sqrt{2} &= \epsilon_1, \\ 2a_0 - 1 - 2a_1\sqrt{2} &= \epsilon_2. \end{aligned}$$

Hence

$$(6) \quad (2a_0 - 1)^2 - 8a_1^2 = \epsilon_1\epsilon_2.$$

Since the left side of (6) is rational it follows that $\epsilon_1\epsilon_2 = 1$. Putting $2a_0 - 1 = x$ and $2a_1 = y$ the assertion follows. Since the general solution of $x^2 - 2y^2 = 1$

is given by

$$x - \sqrt{2}y = (3 - 2\sqrt{2})^p = (1 - \sqrt{2})^{2p},$$

we find that

$$\begin{aligned} x - \sqrt{2}y &\equiv 3^p - 2p \cdot 3^{p-1} \sqrt{2} \\ &\equiv (-1)^p - 2p(-1)^{p-1} \sqrt{2} \pmod{4}. \end{aligned}$$

Thus y is always even, and

$$\frac{1+x}{2} \equiv \frac{1+(-1)^p}{2} \pmod{2},$$

i.e., a_0 is even when p is odd and odd when p is even. Thus the circulants derived from a solution with an even p are congruent to the identity, while those derived from a solution with an odd p are congruent to K .

As the referee pointed out, the two classes of circulants can also be obtained from the fact that every positive definite unimodular 8×8 circulant C is a power of M . For, every power M^n is certainly such a circulant. Conversely, the proof of Lemma 1 shows that there is exactly one such circulant whose characteristic roots are given powers of the characteristic roots of M .

If then n is even, we have $M^n = M^{1/2^n} \cdot M^{1/2^n} \sim I$ and for n odd we have $M^n = M^{1/(n-1)} M M^{1/(n-1)} \sim M$.

REFERENCES

1. H. Minkowski, *Grundlagen für eine Theorie der quadratischen Formen mit ganzzahligen Koeffizienten*, Gesammelte Abhandlungen I (1911), 3-144.
2. L. J. Mordell, *The definite quadratic forms in 8 variables with determinant unity*, J. de Math. pures et appliquées, 17 (1938), 41-46.
3. M. Newman and O. Taussky, *On a generalization of the normal basis in abelian algebraic number fields*, Comm. on Pure and Applied Math. 9 (1956), 85-91.
4. O. Taussky, *Unimodular integral circulants*, Math. Zeitschr. 63 (1955), 286-289.

National Bureau of Standards,
Washington 25, D.C.

SIMULTANEOUS PAIRS OF LINEAR AND QUADRATIC EQUATIONS IN A GALOIS FIELD

ECKFORD COHEN

1. Introduction. Let F denote the Galois field $GF(p^r)$ with p^r elements, where p is an odd prime and r is a positive integer. Suppose further that m and n are arbitrary elements of F and that α_i, β_i ($i = 1, \dots, s$) are *nonzero* elements of F . The purpose of this paper is to evaluate the function $N_s(m, n)$, defined, for an arbitrary positive integer s , to be the number of simultaneous solutions in F of the equations

$$(1.1) \quad \begin{cases} m = \alpha_1 x_1^2 + \dots + \alpha_s x_s^2, \\ n = \beta_1 x_1 + \dots + \beta_s x_s. \end{cases}$$

Explicit formulas for $N_s(m, n)$ are obtained in Theorem 1, and on the basis of this theorem, it is easy to establish the solvability criterion contained in Theorem 2. It follows from the latter criterion that the least value of s for which (1.1) is always solvable is the value $s = 4$. We mention that Theorem 1, in the special case $r = 1$ (that is, in the case of rational congruences (mod p)), reduces to a result of O'Connor and Pall (3; 4) proved by a different method.

It is of interest to compare Dickson's formulas (2, §§64-67) for the number of solutions $N_s(m)$ of the first equation in (1.1) alone, with the results for $N_s(m, n)$ obtained in this paper. As it might be expected, the results for the simultaneous problem are somewhat more involved. A significant difference between the results for the two problems arises from the fact that $N_s(m) > 0$ for all $s \geq 2$.

In this paper we use a direct method based on the trigonometric expansion of $N_s(m, n)$. The most that will be required is a double application of the generalized Cauchy-Gauss sum, (1.7) and (1.11) below.

Next we introduce some notation that will be needed in §2 and §3. Let $t(a)$ denote the trace of an element a in F ,

$$t(a) = a + a^p + \dots + a^{p^{r-1}}.$$

Then we place

$$(1.2) \quad e(a) = e^{2\pi i t(a)/p},$$

from which it follows that $e(a + b) = e(a) e(b)$. The symbol \sum_x will be used to indicate a sum over the totality of elements of F , while $\sum_{x \neq 0}$ will denote a sum over the nonzero elements of F . One will note the property,

$$(1.3) \quad \xi(a) = \sum_x e(ax) = \begin{cases} p^r, & a = 0, \\ 0, & a \neq 0, \end{cases}$$

Received June 4, 1956.

which may be restated in the form,

$$(1.4) \quad c(a) = \sum_{x \neq 0} e(ax) = \begin{cases} p' - 1, & a = 0, \\ -1, & a \neq 0. \end{cases}$$

The symbol $\psi(a)$ will be used to denote the Legendre symbol in F , that is, $\psi(a) = 1, -1$, or 0 according as a is a nonzero square, a non-square, or is zero in F . We denote the quadratic Gauss sums in F by

$$(1.5) \quad G(a) = \sum_x e(ax^2),$$

$$(1.6) \quad G^*(a) = \sum_{x \neq 0} \psi(x) e(ax).$$

The less familiar Cauchy-Gauss sum is defined for F by

$$(1.7) \quad S(a, b) = \sum_x e(ax^2 + 2bx).$$

We mention the following well-known properties of $G(a)$ and $G^*(a)$:

$$(1.8) \quad G(a) = \psi(a) G(1), \quad a \neq 0,$$

$$(1.9) \quad G^2(1) = \psi(-1) p',$$

$$(1.10) \quad G^*(a) = \begin{cases} G(a), & a \neq 0, \\ 0, & a = 0. \end{cases}$$

The sum $S(a, b)$ has the reduction property (1. §6).

$$(1.11) \quad S(a, b) = \begin{cases} e(-b^2/a) G(a), & a \neq 0, \\ p', & a = b = 0, \\ 0, & a = 0, b \neq 0. \end{cases}$$

2. The evaluation of $N_s(m, n)$. We shall need the following additional notation,

$$(2.1) \quad \alpha = \alpha_1 \dots \alpha_s,$$

$$(2.2) \quad \beta = \frac{\beta_1^2}{\alpha_1} + \dots + \frac{\beta_s^2}{\alpha_s},$$

$$(2.3) \quad \gamma = n^2 - \beta m.$$

The results of this section can be stated most conveniently in terms of the five following cases arising from conditions satisfied by m, n, β , and γ .

Case I: $\beta = 0, n \neq 0$,

Case II: $\beta = n = 0, m \neq 0$,

Case III: $\beta = m = n = 0$,

Case IV: $\beta \neq 0, \gamma \neq 0$,

Case V: $\beta \neq 0, \gamma = 0$.

We now prove

THEOREM 1. *The number of solutions $N_s(m, n)$ of (1.1) is given by*

$$(2.4) \quad N_s(m, n) = \begin{cases} p^{r(4k-2)} + p^{r(2k-1)} \psi(\alpha) \zeta, & s = 4k, \\ p^{r(4k-1)} + p^{r(2k-1)} \psi(\alpha) \eta, & s = 4k+1, \\ p^{4kr} + p^{2kr} \psi(-\alpha) \zeta, & s = 4k+2, \\ p^{r(4k+1)} + p^{2kr} \psi(-\alpha) \eta, & s = 4k+3, \end{cases}$$

where η and ζ are defined by $\eta = 0, \zeta = 0$ in Case I; $\eta = p^r \psi(m), \zeta = -1$ in Case II; $\eta = 0, \zeta = p^r - 1$ in Case III; $\eta = -\psi(\beta), \zeta = \psi(\gamma)$ in Case IV; $\eta = (p^r - 1) \psi(\beta), \zeta = 0$ in Case V.

Remark. It is to be understood that $N_s(m, n)$ is undefined for any cases that may be incompatible.

Proof. The function $N_s(m, n)$ has the double Fourier expansion (5),

$$(2.5) \quad N_s(m, n) = p^{-2r} \sum_u \sum_v A(u, v) e(-mu) e(-2nv),$$

$$A(u, v) = \sum_{x_1, \dots, x_s} e\left(u(\alpha_1 x_1^2 + \dots + \alpha_s x_s^2)\right) e\left(2v(\beta_1 x_1 + \dots + \beta_s x_s)\right).$$

We break up this expansion into two parts according as $u = 0$ or $u \neq 0$, to get

$$(2.6) \quad N_s(m, n) = \Sigma_1 + \Sigma_2,$$

where

$$(2.7) \quad \Sigma_1 = p^{-2r} \sum_v e(-2nv) \prod_{i=1}^s \xi(2\beta_i v),$$

$$(2.8) \quad \Sigma_2 = p^{-2r} \sum_{u \neq 0} \sum_v e(-mu) e(-2nv) \prod_{i=1}^s S(\alpha_i u, \beta_i v).$$

By (1.3) we have immediately

$$(2.9) \quad \Sigma_1 = p^{r(s-2)}.$$

Now by (1.8) and (1.11) one obtains for $u \neq 0$,

$$S(\alpha_i u, \beta_i v) = e\left(\frac{-\beta_i^{2r^2}}{\alpha_i u}\right) \psi(\alpha_i u) G(1),$$

so that (2.8) becomes, using the definition of β ,

$$(2.10) \quad \Sigma_2 = G^s(1) p^{-2r} \psi(\alpha) \sum_{u \neq 0} \psi^s(u) e(-mu) S(-\beta/u, -n).$$

If $u \neq 0$, we have, again by (1.8) and (1.11),

$$(2.11) \quad S(-\beta/u, -n) = \begin{cases} e(n^2 u / \beta^s) \psi(-\beta u) G(1), & \beta \neq 0, \\ p^r, & \beta = n = 0, \\ 0, & \beta = 0, n \neq 0. \end{cases}$$

We now evaluate Σ_2 in the separate cases arising from (2.11). It follows immediately from (2.10) that

$$(2.12) \quad \Sigma_2 = 0, \quad \beta = 0, n \neq 0.$$

In case $\beta = n = 0$, we obtain from (2.10) and (2.11),

$$(2.13) \quad \sum_2 = \begin{cases} G^s(1) p^{-r} \psi(\alpha) c(-m), & \beta = n = 0, s \text{ even}, \\ G^s(1) p^{-r} \psi(\alpha) G^s(-m), & \beta = n = 0, s \text{ odd}. \end{cases}$$

Applying (1.4), (1.9), and (1.10) to (2.13), it follows, in case s is even, that

$$(2.14) \quad \sum_2 = \begin{cases} -\psi((-1)^{1/2} \alpha) p^{1/2r(s-2)}, & \beta = n = 0, m \neq 0, s \text{ even}, \\ \psi((-1)^{1/2} \alpha) p^{1/2r(s-2)} (p^r - 1), & \beta = m = n = 0, s \text{ even}, \end{cases}$$

and in case s is odd,

$$(2.15) \quad \sum_2 = \begin{cases} \psi((-1)^{1/2(s+1)} \alpha m) p^{1/2r(s-1)}, & \beta = n = 0, m \neq 0, s \text{ odd}, \\ 0, & \beta = m = n = 0, s \text{ odd}. \end{cases}$$

In case $\beta \neq 0$, it follows from (2.10) and (2.11) that

$$(2.16) \quad \sum_2 = \begin{cases} G^{s+1}(1) p^{-2r} \psi(-\alpha\beta) G^s(\gamma/\beta), & \beta \neq 0, s \text{ even}, \\ G^{s+1}(1) p^{-2r} \psi(-\alpha\beta) c(\gamma/\beta), & \beta \neq 0, s \text{ odd}. \end{cases}$$

Applying (1.4), (1.9), and (1.10) to (2.16), we obtain, in case s is even,

$$(2.17) \quad \sum_2 = \begin{cases} \psi((-1)^{1/2(s+1)} \alpha\gamma) p^{1/2r(s-2)}, & \beta \neq 0, \gamma \neq 0, s \text{ even}, \\ 0, & \beta \neq 0, \gamma = 0, s \text{ even}, \end{cases}$$

and in case s is odd,

$$(2.18) \quad \sum_2 = \begin{cases} -\psi((-1)^{1/2(s+1)} \alpha\beta) p^{1/2r(s-3)}, & \beta \neq 0, \gamma \neq 0, s \text{ odd}, \\ \psi((-1)^{1/2(s+1)} \alpha\beta) p^{1/2r(s-3)} (p^r - 1), & \beta \neq 0, \gamma = 0, s \text{ odd}. \end{cases}$$

Combining (2.6), (2.9), (2.12), (2.14), (2.15), (2.17), and (2.18) the theorem follows.

3. Solvability criterion. We now apply Theorem 1 to the cases $s < 4$ to obtain the following explicit results.

$$(3.1) \quad N_1(m, n) = \begin{cases} 1, & \text{Case V,} \\ 0, & \text{Case IV;} \end{cases}$$

$$(3.2) \quad N_2(m, n) = \begin{cases} 1, & \text{Cases I, V,} \\ 0, & \text{Case II,} \\ p^r, & \text{Case III,} \\ 1 + \psi(-\alpha\gamma), & \text{Case IV;} \end{cases}$$

$$(3.3) \quad N_2(m, n) = \begin{cases} p^r, & \text{Cases I, III,} \\ p^r + p^r \psi(-\alpha m), & \text{Case II,} \\ p^r - \psi(-\alpha \beta), & \text{Case IV,} \\ p^r + (p^r - 1) \psi(-\alpha \beta), & \text{Case V;} \end{cases}$$

$$(3.4) \quad N_4(m, n) = \begin{cases} p^{2r}, & \text{Cases I, V,} \\ p^{2r} - p^r \psi(\alpha), & \text{Case II,} \\ p^{2r} + p^r (p^r - 1) \psi(\alpha), & \text{Case III,} \\ p^{2r} + p^r \psi(\alpha \gamma), & \text{Case IV.} \end{cases}$$

It is noted that Cases I, II, and III do not arise if $s = 1$ or if $s = 2$ and $\psi(-\alpha) = -1$.

On the basis of (3.1), (3.2), (3.3), and (3.4) we obtain immediately the following solvability criterion.

THEOREM 2. *Subject to the restrictions stated in the Introduction, (1.1) is always solvable ($N_s(m, n) > 0$) provided $s \geq 4$. The only cases in which (1.1) is insolvable, that is when $N_s(m, n) = 0$, are the following:*

- (1) $s = 1, \quad \gamma \neq 0,$
- (2) $s = 2, \quad \beta \neq 0, \quad \gamma \neq 0, \quad \psi(-\alpha \gamma) = -1,$
- (3) $s = 2, \quad \beta = n = 0, \quad m \neq 0,$
- (4) $s = 3, \quad \beta = n = 0, \quad m \neq 0, \quad \psi(-\alpha m) = -1,$

where α, β , and γ are defined as in §2.

REFERENCES

1. Leonard Carlitz, *Weighted quadratic partitions over a finite field*, Can. J. Math., 5 (1953), 317-323.
2. L. E. Dickson, *Linear Groups* (Leipzig, 1901).
3. R. E. O'Connor, *Quadratic and linear congruence*, Bull. Amer. Math. Soc., 45 (1939), 792-798.
4. R. E. O'Connor and Gordon Pall, *The quaternion congruence $ia \equiv b \pmod{g}$* , Amer. J. Math., 61 (1939), 487-508.
5. A. L. Whiteman, *Finite Fourier series and equations in finite fields*, Trans. Amer. Math. Soc., 54 (1953), 78-98.

University of Tennessee

THE SET OF ALL GENERALIZED LIMITS OF BOUNDED SEQUENCES

MEYER JERISON

1. Introduction. Let M be the normed linear space whose general element, x , is a bounded sequence

$$\{\xi_n\}_{n=1}^{\infty}$$

of real numbers, and $\|x\| = \text{l.u.b. } |\xi_n|$. Let T denote the linear operation (of norm 1) defined by $Tx = (\xi_2, \xi_3, \dots, \xi_{n+1}, \dots)$. A *generalized limit* is a linear functional ϕ on M which satisfies the conditions

- (1) $x \geq 0$ (i.e., $\xi_n \geq 0$ for all n) implies $\phi(x) \geq 0$;
- (2) $\phi(Tx) = \phi(x)$ for all $x \in M$;
- (3) $\phi(1, 1, 1, \dots) = 1$.

The set of all generalized limits will be denoted by L . In the presence of (1), condition (3) is equivalent to $\|\phi\| = 1$.

The basic question of existence of generalized limits has been settled in a variety of ways; the standard proof appears in (2, p. 34). This proof, based upon the Hahn-Banach theorem, actually leads to all generalized limits, and this fact was used in (7) to obtain properties of L . In the present paper, attention is focused on another existence proof (5, p. 1010; 8; 10, p. 52) which depends, ultimately, on Tychonoff's theorem. In order to describe this proof, we must summarize some well-known properties of M and of the conjugate space M^* (1; 5).

Only one topology in M^* will be of interest to us, namely, the weak* topology, which is defined as follows: A directed system $\{\phi_\alpha\}$ in M^* converges to ϕ if $\phi_\alpha(x) \rightarrow \phi(x)$ for each $x \in M$. An essential property of this topology is that the set

$$B^* = \{\phi \mid |\phi(x)| \leq \|x\| \text{ for all } x \in M\}$$

(the unit ball) is compact. B^* is also convex, and we are able to apply the Krein-Milman theorem to its subsets. Thus, if K is a closed convex subset of B^* , and $S (\subset K)$ contains all extreme points of K , then K is the closed convex hull of S , denoted by \bar{S} . In particular, if K is not empty, neither is S .

We will denote by Ω' the set of extreme points of B^* that satisfy condition (1), or equivalently, the collection of extreme points of the subset of B^* which is determined by conditions (1) and (3). Since the latter set is closed and convex, it is, in fact, $\bar{\Omega}'$. Among the functionals in Ω' are those of the form $\theta_p(x) = \xi_p$ for each fixed natural number p . The collection N of all such

Received October 26, 1953; in revised form October 21, 1955. This work was supported, in part, by a grant from the National Science Foundation (USA).

functionals is a discrete, open, and dense subset of Ω' . Since Ω' happens to be by compact (1, p. 504), it is the closure of N in M^* , and it is known as the Stone-Čech compactification of N . The set Ω , the complement of N in Ω' , is characterized, as a subset of Ω' , by the fact that each of its elements satisfies the condition

(4) $\phi(x)$ is independent of the value of ξ_n for each fixed n .

Its closed convex hull, $\mathfrak{S}(\Omega)$, is the set of all functionals in M^* satisfying conditions (1), (3), and (4). Since (2) obviously implies (4), $L \subset \mathfrak{S}(\Omega)$. It is important to note that L is itself a closed convex set.

The proof of existence of generalized limits that was referred to in the second paragraph goes as follows: Let $\chi \in \Omega$, or, more generally, $\chi \in \mathfrak{S}(\Omega)$; a generalized limit, ψ , is obtained by setting

$$\psi(x) = \chi\left(\left\{n^{-1} \sum_{i=1}^n \xi_i\right\}\right)$$

for $x = \{\xi_n\}$ in M . Now, a functional $\chi \in \mathfrak{S}(\Omega)$ has the property that if $y \in M$ is a convergent sequence, then $\chi(y)$ is the ordinary limit of y . If x is a sequence whose arithmetic means, $n^{-1} \sum \xi_i$, converge to σ , say, then $\psi(x) = \sigma$ for every generalized limit ψ which is obtained in this manner. Since it is known that there are some such sequences and some generalized limits which assign to them values different from σ (see Theorem 4), it follows that this procedure does not lead to all generalized limits. It is our intention, therefore, to modify this procedure in such a way as to obtain all generalized limits.

Let T_n be the operator

$$n^{-1} \sum_{i=0}^{n-1} T^i,$$

and let $\theta_1(x) = \xi_1$ for all $x \in M$. Then $\{\theta_1(T_n x)\}$ is the sequence of arithmetic means of x , and the generalized limit obtained above may be defined as $\psi(x) = \chi(\{\theta_1(T_n x)\})$. From this point of view, an obvious way to generate more generalized limits is to replace θ_1 by some other functional. Three observations should be made in this connection.

(a) If θ_1 is replaced by $\theta_p \in N$, nothing new is obtained, because

$$\theta_p(T^i x) = \theta_1(T^{i+p-1} x), \quad i = 1, 2, \dots,$$

so that

$$\chi(\{\theta_p(T_n x)\}) = \chi(\{\theta_1(T_n T^{p-1} x)\}) = \psi(T^{p-1} x) = \psi(x).$$

(b) If ϕ is any functional satisfying conditions (1) and (3), i.e., $\phi \in \mathfrak{S}(\Omega')$, and if $\chi \in \mathfrak{S}(\Omega)$, then

$$(5) \quad \psi(x) = \chi(\{\phi(T_n x)\})$$

does define a generalized limit (Theorem 1).

(c) It is trivial that any generalized limit ψ may be obtained from (5) simply by taking $\phi = \psi$, because $\psi(T_n x) = \psi(x)$ for all n .

The purpose of this paper is to prove that the collection Q of all functionals of the form (5) with χ and ϕ in Ω is sufficient to yield all generalized limits in the sense that $\mathfrak{S}(Q) = L$.

The proof requires a surprising amount of heavy machinery: the representation of M as the space of all continuous functions on Ω' , the representation of an arbitrary continuous linear functional on M as a measure on Ω' , and one of the deep theorems of measure theory, the individual ergodic theorem. A similar (but apparently weaker) result can be obtained by using entirely different techniques. As was stated before, the set N is dense in Ω' , which means that the element $\chi \in \Omega$ is the limit of a directed system

$$\{\theta_{n_r}\}$$

of elements of N . In terms of this directed system, (5) becomes

$$\psi(x) = \lim_r \phi(T_{n_r} x).$$

Since $\chi \in \Omega$, $n_r \rightarrow \infty$. Let us consider, now, the set A of all continuous linear functionals of the form

$$(6) \quad \psi(x) = \lim_r \phi_r(T_{n_r} x)$$

where $n_r \rightarrow \infty$, $\phi_r \in \Omega$ for all r , and the limit is assumed to exist for all $x \in M$. It will be seen that A is a closed subset of L , and the proposition $\mathfrak{S}(A) = L$ will be proved independently of the obvious fact that $Q \subset A$. I have not been able to determine¹ whether Q is a proper subset of A or whether the closure of Q is all of A .

2. Limit points of sequences in M^* . Since it is convenient to work in the space M^* as much as possible, we introduce the operator T^* on M^* , defined by $T^* \phi(x) = \phi(Tx)$ for all $x \in M$, $\phi \in M^*$. Condition (2) becomes (2*) $T^* \phi = \phi$. In keeping with the notation used above, we have

$$T_n^* = n^{-1} \sum_{i=0}^{n-1} T^{*i}.$$

THEOREM 1. *If*

$$\psi = \lim_r T_{n_r}^* \phi_r,$$

where each ϕ_r is a positive linear functional of norm 1 (i.e., ϕ_r satisfies conditions (1) and (3)) and $n_r \rightarrow \infty$, then $\psi \in L$.

Proof. It is clear that ψ satisfies conditions (1) and (3). For any $x \in M$,

¹I am grateful to the referee for many suggestions for the improvement of an earlier version of this paper. Most important, by far, was his discovery of an error in what purported to be a proof that $Q = A$.

$$\begin{aligned}\psi(x - Tx) &= \lim_p [T_{n_p}^* \phi_p(x) - T_{n_p}^* \phi_p(Tx)] \\ &= \lim_p \phi_p \left[n_p^{-1} \sum_{i=0}^{n_p-1} (T^i x - T^{i+1} x) \right] \\ &= \lim_p \phi_p [n_p^{-1} (x - T^{n_p} x)].\end{aligned}$$

But

$$|\phi_p [n_p^{-1} (x - T^{n_p} x)]| \leq 2 \|\phi_p\| \cdot n_p^{-1} \cdot \|x\| \rightarrow 0$$

with p . Therefore, $\psi(x - Tx) = 0$.

The existence of generalized limits is an immediate consequence of Theorem 1; this is substantially the same proof that was described in the Introduction. Take any $\theta \in N$. Since $\{T_n^* \theta\} \subset B^*$ and the latter set is compact, the former set must have limit points all of which are in L , according to Theorem 1. We observe that $\{T_n^* \theta\}$ is not a convergent sequence in M^* for any $\theta \in N$, because convergence of $\{T_n^* \theta\}$ in M^* implies convergence of the sequence of numbers $\{T_n^* \theta(x)\}$ for every $x \in M$. But, if $\theta = \theta_p$, then

$$T_n^* \theta_p(x) = \theta_p(T_n x)$$

is the n th arithmetic mean of the sequence $T^{p-1} x$, and it is easy to find an x which will make $\{\theta_p(T_n x)\}$ diverge. What is not so obvious is that if $\omega \in \Omega$, $\{T_n^* \omega\}$ need not converge either. Here is an example.

First, construct a sequence $\{\eta_i\}$ whose arithmetic means do not converge. Thus: $\eta_1 = 1$, $\eta_i = 0$ if $2^k \leq i < 2^k + 2^{k-1}$, and $\eta_i = 1$ if

$$2^k + 2^{k-1} \leq i < 2^{k+1}, \quad k = 1, 2, \dots$$

Next, let

$$\{\theta_{n_k}\} (k = 1, 2, \dots)$$

be a sequence in N such that $n_k - n_{k-1} \geq 2^k$, and let ω be one of its limit points. Define $\xi_n = 0$ if $n = n_k$ for some k or $n < n_1$, and $\xi_n = \eta_i$ if i is the least positive integer such that $n = n_k + i$ for some k . The inequality serves to guarantee that, for each i ,

$$\xi_{n_k+i} = \eta_i$$

for all but a finite number of k 's. Setting $x = \{\xi_n\}$, we have

$$\omega(x) = \lim_{k \rightarrow \infty} \xi_{n_k}$$

provided the limit exists, and more generally,

$$\omega(T^i x) = \lim_{k \rightarrow \infty} \xi_{n_k+i}.$$

It follows that for each i , $\omega(T^i x) = \eta_i$, and therefore,

$$\left\{ n^{-1} \sum_{i=0}^{n-1} \omega(T^i x) \right\}$$

does not converge. But this implies that $\{T_n^* \omega\}$ does not converge in M^* .

The sequence $\{n_k\}$ is an example of a sequence of integers of density 0 in the number theoretic sense. Making use of the obvious one to one correspondence between N and the set of positive integers, we may speak, in the same way, of non-dense subsets of N . It is somewhat surprising, in view of the fact that N is a countable, dense, discrete set in Ω' , that there exist points ω in Ω which are not limit points of any non-dense subset of N . I conjecture that for some such ω the sequence $\{T_n^* \omega\}$ does converge, but I have not been able to prove this. It will be seen in the proof of Theorem 3 that for each $x \in M$ there exists $\omega \in \Omega$ with the property that $\{T_n^* \omega(x)\}$ is a convergent sequence of numbers.

3. Sets that generate L . We have already observed that $L \subset \mathfrak{S}(\Omega)$. Since T_n^* is a continuous mapping and Ω is compact, the set $A_n = T_n^*(\Omega)$ is compact. From the additional fact that T_n^* is linear, it follows that for any set $S \subset B^*$,

$$T_n^*(\mathfrak{S}(S)) = \mathfrak{S}(T_n^*(S)).$$

Since L is elementwise invariant under T_n^* , we have

$$L = T_n^*(L) \subset T_n^*(\mathfrak{S}(\Omega)) = \mathfrak{S}(T_n^*(\Omega)) = \mathfrak{S}(A_n).$$

It follows that

$$L \subset \bigcap_1^\infty \mathfrak{S}(A_n).$$

This conclusion is obtained so easily because the closed convex hull of a set is, in general, very much bigger than the set itself. We would like to have L as the closed convex hull of an intersection rather than the intersection of hulls. It is hopeless, however, to expect that

$$L = \mathfrak{S}\left(\bigcap_1^\infty A_n\right),$$

because $\bigcap A_n$ is empty in the worst possible way; namely, the sets A_n are mutually disjoint. To prove the last statement, fix n , and let $x = \{\xi_i\}$ where $\xi_i = 1$ if i is a multiple of n and $\xi_i = 0$ otherwise. Then for every $\omega \in \Omega'$, $\omega(x)$ is either 0 or 1, there is some integer j' such that $T^{*j'} \omega(x) = 1$, and $T^{*j} \omega(x) = 1$ if $j \equiv j' \pmod{n}$ and $T^{*j} \omega(x) = 0$ if $j \not\equiv j' \pmod{n}$. Consequently, $\phi(x) = 1/n$ for all $\phi \in A_n$, whereas, if $\phi \in A_k$ with $k < n$, then $\phi(x) = 0$ or $1/k$. Thus, $A_k \cap A_n$ is empty for all $k < n$ and for all n .

The operation on the sequence of sets $\{A_n\}$ that does the job we want is the topological limit superior. By definition,

$$\psi \in \limsup A_n$$

if every neighborhood of ψ meets infinitely many of the sets A_n . Equivalently,

$$\limsup A_n = \bigcap_{m=1}^\infty F_m,$$

where F_n is the closure of the set

$$\bigcup_{n=m}^{\infty} A_n.$$

It is clear from the form of the sets A_n that if $\psi \in \limsup A_n$, then there are directed systems

$$\{T_n, *\} \text{ and } \{\omega_n\}$$

directed by the same set $\{\nu\}$ and with $\omega_n \in \Omega$, such that $n_r \rightarrow \infty$ and

$$T_{n_r}^* \omega_{n_r} \rightarrow \psi.$$

Conversely, every cluster point of such a directed system is in $\limsup A_n$. Thus, $\limsup A_n$ is precisely the set A of the introduction. By Theorem 1, $A \subset L$, and since L is compact and convex, we conclude that $\mathfrak{S}(A) \subset L$.

THEOREM 2. $\mathfrak{S}(A) = L$.

Proof. In view of the preceding discussion, we have only to prove that $L \subset \mathfrak{S}(A)$. But this follows easily from Theorem 2 of (4), which asserts that for a sequence $\{\mathfrak{S}(A_n)\}$ of compact convex sets,

$$\limsup \mathfrak{S}(A_n) \subset \mathfrak{S}(\limsup A_n).$$

We have already seen that

$$L \subset \bigcap_1^{\infty} \mathfrak{S}(A_n),$$

which, in turn, is contained in $\limsup \mathfrak{S}(A_n)$.

COROLLARY 1. All of the extreme points of L are in A .

Proof. The topological limit superior is always a closed set. According to a theorem of Milman (9; 3, p. 84, prop. 4), all of the extreme points of the closed convex hull of a set are in the closure of that set.

COROLLARY 2.

$$L = \bigcap_{n=1}^{\infty} T_n^*(\mathfrak{S}(\Omega)).$$

This can be proved directly, but at this stage it comes very easily out of a string of inequalities:

$$\begin{aligned} L &= \bigcap_{n=1}^{\infty} T_n^*(L) \subset \bigcap_{n=1}^{\infty} T_n^*(\mathfrak{S}(\Omega)) = \bigcap_{n=1}^{\infty} \mathfrak{S}(A_n) \\ &\subset \limsup \mathfrak{S}(A_n) \subset \mathfrak{S}(\limsup A_n) = L. \end{aligned}$$

We turn now to the development of a sharper result. It is well known that M is isomorphic and isometric (equivalent in Banach's sense) with the space of all continuous functions on the compact set Ω' , the isomorphism taking the element $x \in M$ into the function whose value at $\omega \in \Omega'$ is $\omega(x)$. It is convenient

to identify the two spaces and use the notation $x(\omega)$ to denote $\omega(x)$. To each continuous linear functional ϕ on M , there corresponds (5) a unique regular Borel measure m on Ω' with the property

$$\phi(x) = \int_{\Omega'} x(\omega) dm(\omega) \quad \text{for all } x \in M.$$

Conditions (1)–(4) on the functional ϕ may be translated into the following conditions on the measure m : For every Borel set Γ ,

$$(1') \quad m(\Gamma) \geq 0;$$

$$(2') \quad m(T^* \Gamma) = m(\Gamma);$$

$$(3') \quad m(\Omega') = 1;$$

$$(4') \quad m(N) = 0.$$

THEOREM 3. *If Q is the set of all limit points of sequences $\{T_n^* \omega\}$, $\omega \in \Omega$, then $\mathfrak{S}(Q) = L$.*

Proof. We apply here the technique of Milman, as described in (4). Since L is a compact, convex set and $Q \subset L$, it suffices to prove that for every continuous (in the weak* topology) linear functional f on M^* ,

$$\sup_{\psi \in L} f(\psi) = \sup_{\psi \in Q} f(\psi).$$

But every continuous linear functional on M^* comes from an element of M ; that is, given such an f , there exists $x \in M$ such that $f(\phi) = \phi(x)$ for all $\phi \in M^*$. We wish to prove, then, that for all $x \in M$,

$$(7) \quad \sup_{\psi \in L} \psi(x) = \sup_{\psi \in Q} \psi(x).$$

Fix $x \in M$. Since L is compact, there is an element ψ_0 in L such that

$$\psi_0(x) = \sup_{\psi \in L} \psi(x).$$

Let m_0 be the regular Borel measure on Ω' corresponding to the functional ψ_0 . Conditions (1')–(4') are satisfied by m_0 , so that all of the measure is carried by the set Ω : $m_0(\Omega) = 1$. The individual ergodic theorem (6) is applicable to this situation, and we conclude that $\lim_{n \rightarrow \infty} T_n x(\omega)$ exists for all ω in an invariant (under T^*) subset Δ of Ω of m_0 -measure 1. If we let $X(\omega)$ denote this limit for $\omega \in \Delta$, and $X(\omega) = 0$ for $\omega \in \Omega - \Delta$, then the ergodic theorem states further that X is a measurable function, invariant under T^* , and

$$(8) \quad \int_{\Omega} X(\omega) dm_0(\omega) = \int_{\Omega} x(\omega) dm_0 = \psi_0(x).$$

For each $\omega \in \Delta$, the set $\{T_n^* \omega\}$ has limit points, and according to Theorem 1, every such limit point is in L . Consequently, there is a ψ in L , which depends upon ω and such that

$$\lim_{n \rightarrow \infty} T_n x(\omega) = \psi(x),$$

Since $\psi(x) \leq \psi_0(x)$ for all $\psi \in L$, we have

$$X(\omega) = \lim T_n x(\omega) \leq \psi_0(x)$$

for all $\omega \in \Delta$. This inequality together with equation (8) and the fact that $m_0(\Delta) = 1$ implies that $X(\omega) = \psi_0(x)$ for almost all $\omega \in \Delta$. In particular, there is at least one ω_0 for which the equality holds, and any limit point χ of $\{T_n^* \omega_0\}$ has the properties:

$$\chi \in Q \text{ and } \chi(x) = \psi_0(x) = \sup_{\psi \in L} \psi(x).$$

This verifies (7) and completes the proof of the theorem (cf. 6a).

4. Almost convergent sequences. Lorentz (7) calls a sequence $x \in M$ *almost convergent* if $\psi(x)$ is independent of $\psi \in L$. By Theorem 3, it is obviously sufficient to require that $\psi(x)$ be constant for $\psi \in Q$. Observing that every $\psi \in Q$ has the form

$$\psi(x) = \lim_n T_{n_k} x(\omega), \quad \omega \in \Omega,$$

leads to the following characterization of almost convergent sequences:

LEMMA. *In order that there exist a number σ such that $\psi(x) = \sigma$ for all $\psi \in L$, it is necessary and sufficient that*

$$(9) \quad \lim_{n \rightarrow \infty} T_n x(\omega) = \sigma$$

for all $\omega \in \Omega'$.

(We write Ω' rather than Ω in order to facilitate the coming discussion. Both are correct.)

The known characterization of almost convergent sequences is the following (7; see also 10, p. 53):

THEOREM 4. *A necessary and sufficient condition for the existence of σ such that $\psi(x) = \sigma$ for all $\psi \in L$ is that*

$$(10) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \xi_{i+k} = \sigma,$$

uniformly in k .

In our notation, (10) takes the form

$$(10') \quad \lim_{n \rightarrow \infty} T_n x(k) = \sigma,$$

uniformly in $k \in N$. Since N is dense in Ω' and $T_n x$ is a continuous function on Ω' , (10') obviously implies (9). But the converse is also true. For (9) implies weak convergence of the sequence $\{T_n x\}$ in M to the constant σ . By the mean ergodic theorem in Banach space (6), this implies convergence in norm, i.e. uniform convergence on all of Ω' .

5. The maximal generalized limit of a given sequence. If ρ is a functional on M satisfying the two conditions

- (i) $\rho(x + y) \leq \rho(x) + \rho(y)$,
- (ii) $\rho(\lambda x) = \lambda \rho(x)$ for all $x, y \in M$ and $\lambda \geq 0$,

then, according to the Hahn-Banach theorem, there exists a linear functional ϕ such that $\phi(x) \leq \rho(x)$ for all x . The functional τ defined by

$$\tau(x) = \sup_{\phi \in L} \phi(x)$$

satisfies conditions (i) and (ii). Noting that for a linear functional ϕ , $\phi(x) \leq \tau(x)$ for all $x \in M$ implies

$$-\tau(-x) \leq -\phi(-x) = \phi(x) \leq \tau(x),$$

one sees that every such ϕ is actually a generalized limit.

It is clear from the proof of the Hahn-Banach theorem (2, p. 28) that not only is there a linear functional ϕ dominated by ρ , but also, if $x_0 \in M$ is given, ϕ may be chosen so that $\phi(x_0) = \rho(x_0)$. Consequently, if ρ has the property that $\phi \leq \rho$ implies $\phi \in L$, then $\rho(x) \leq \tau(x)$ for all $x \in M$. In Banach's proof of the existence of generalized limits the functional which is used for ρ is

$$\tau'(x) = \inf_{k \rightarrow \infty} \limsup n^{-1} \sum_{i=1}^n \xi_{k+m_i},$$

where the infimum is taken over all possible choices of non-negative integers m_1, \dots, m_n . By what was said earlier, $\tau'(x) \leq \tau(x)$, and it is easy to reverse this inequality to obtain $\tau'(x) = \tau(x)$. All of this has been observed before (7).

Now, let us use Theorem 2 to calculate $\tau(x)$ from the terms of the sequence $\{\xi_n\}$. Since $L = \mathfrak{S}(A)$, where $A = \limsup A_n$, we have

$$\tau(x) = \sup_{\phi \in A} \phi(x).$$

By definition,

$$A = \bigcap_{m=1}^{\infty} F_m,$$

where F_m is the closure of

$$\bigcup_{n=m}^{\infty} A_n.$$

It is easy to see that

$$\sup_{\phi \in A_n} \phi(x) = \limsup_{k \rightarrow \infty} n^{-1} \sum_{i=0}^{n-1} \xi_{k+i},$$

so that

$$\sup_{\phi \in F_m} \phi(x) = \sup_{n \geq m} \limsup_{k \rightarrow \infty} n^{-1} \sum_{i=0}^{n-1} \xi_{k+i},$$

and finally,

$$(11) \quad \sup_{\phi \in A} \psi(x) \leq \inf_m \sup_{n \geq m} \limsup_{k \rightarrow \infty} n^{-1} \sum_{i=0}^{n-1} \xi_{k+i}.$$

We denote the right side of the inequality by $\tau''(x)$:

$$\tau''(x) = \limsup_{n \rightarrow \infty} \limsup_{k \rightarrow \infty} n^{-1} \sum_{i=0}^{n-1} \xi_{k+i}.$$

Since F_m is a compact set, the supremum of $\phi(x)$ is attained at some $\phi_m \in F_m$, for each m . The sequence $\{\phi_m\}$ has limit points and every one of these limit points is in A . It follows that the inequality in (11) is actually an equality.

THEOREM 5. For every $x = \{\xi_n\}$ in M ,

$$\sup_{\psi \in L} \psi(x) = \tau'(x) = \tau''(x) = \lim_{n \rightarrow \infty} \limsup_{k \rightarrow \infty} n^{-1} \sum_{i=0}^{n-1} \xi_{k+i}.$$

Everything has already been proved excepting the existence of the last limit, which is obtained by observing that

$$\tau'(x) \leq \liminf_{n \rightarrow \infty} \limsup_{k \rightarrow \infty} n^{-1} \sum_{i=0}^{n-1} \xi_{k+i} \leq \tau''(x).$$

An alternate proof of the equality $\tau' = \tau$ can be given by using the Hahn-Banach argument mentioned earlier; namely, if ϕ is a linear functional such that $\phi(x) \leq \tau''(x)$ for all x , then $\phi \in L$. Consequently, $\tau' \leq \tau$, and a simple computation shows that $\psi \in L$ implies $\psi \leq \tau''$. The argument based upon Theorem 2, although more complicated than this one, has the advantage that it enables one to discover the form of the functional τ'' . In fact, the equality $\tau' = \tau''$ does not seem to have been noticed before.

It ought to be possible to prove that $\tau' = \tau''$ directly from the expressions for τ' and τ'' in terms of the sequence $\{\xi_n\}$, and this was done by Professor J. H. B. Kemperman. The essential step in his proof (unpublished) is the following inequality: if

$$0 = m_1 < m_2 < \dots < m_n,$$

and p is any positive integer, then

$$\limsup_{k \rightarrow \infty} p^{-1} \sum_{i=1}^p \xi_{k+i} \leq \limsup_{k \rightarrow \infty} n^{-1} \sum_{i=1}^n \xi_{k+m_i} + \frac{2\|x\|}{np} (n-1)(m_n - m_1).$$

REFERENCES

1. R. F. Arens and J. L. Kelly, *Characterizations of the space of continuous functions over a compact Hausdorff space*, Trans. Amer. Math. Soc., **62** (1947), 499-508.
2. S. Banach, *Théorie des opérations linéaires* (Warsaw 1932).
3. N. Bourbaki, *Espaces vectoriels topologiques* (Actualités scientifiques et industrielles, no. 1180, Paris, 1953).
4. M. Jerison, *A property of extreme points of compact convex sets*, Proc. Amer. Math. Soc., **5** (1954), 782-783.
5. S. Kakutani, *Concrete representations of abstract (M)-spaces*, Ann. of Math., **42** (1941), 994-1024.
6. S. Kakutani, *Ergodic theory*, Proc. Internat. Congr. Math., Cambridge, 1950, II, 128-142.
- 6a. N. Kryloff and N. Bogoliouboff, *La théorie générale de la mesure dans son application à l'étude des systèmes dynamiques de la mécanique non linéaire*, Ann. of Math., **38** (1937), 65-113.
7. G. G. Lorentz, *A contribution to the theory of divergent sequences*, Acta Math., **80** (1948), 167-190.
8. S. Mazur, *On the generalized limit of bounded sequences*, Colloq. Math., **2** (1951), 173-175.
9. D. Milman, *Characteristics of extremal points of regularly convex sets*, Doklady Akad. Nauk SSSR (N.S.), **57** (1947), 119-122.
10. J. von Neumann, *Invariant Measures* (Institute for Advanced Study, Princeton, 1940-1941).

Purdue University

A NOTE ON ASYMPTOTIC SERIES

HARRY F. DAVIS

Introduction. We extend some observations of Popken (2) on the algebraic foundations of the theory of asymptotic series. The main result is the theorem in §5 which characterizes, for a particular function space, a class of linear functionals defined in §4. In §3 we discuss another class of linear functionals related to asymptotic series. In the first two paragraphs we give definitions which render this note self-contained.

This note grew out of a department seminar led by T. E. Hull, to whom I am indebted for many stimulating discussions.

1. Asymptotic sums. Let $c_0 + c_1x + c_2x^2 + \dots$ be a formal power series with real coefficients. By an *asymptotic sum* (as $x \rightarrow 0+$) of the series, is meant a real-valued function f , having as domain the positive reals, such that for all nonnegative integers n ,

$$f(x) - c_0 - c_1x - \dots - c_nx^n = o(x^n), \quad x \rightarrow 0+.$$

If f has this property, so also has, for example, $f(x) + e^{-1/x}$; the asymptotic sum of a formal series is therefore not unique. It is known that every such series has an asymptotic sum, which can be constructed in the following way. Let S denote a nested sequence

$$N_0 \supset N_1 \supset N_2 \supset \dots$$

of neighborhoods of $x = 0$ with characteristic functions u_0, u_1, u_2, \dots . Assume further that any positive x is in one and at most finitely many of these neighborhoods, so the series

$$c_0u_0(x) + c_1u_1(x) + c_2u_2(x)x^2 + \dots$$

converges for all positive x and defines a function $f(S, x)$. Then for a suitable nest S , whose choice depends on the power series, the function $f(S, x)$ is an asymptotic sum of the formal series. If we select a fixed S , and associate with each formal power series the corresponding $f(S, x)$, we obtain an isomorphism between the ring of all formal power series and a ring of functions. Since it is impossible to find a nest S that gives an asymptotic sum for all series, this isomorphism pairs each series with an asymptotic sum of the series only over a subring. We do not know if, by some other method, it is possible to obtain a correspondence between the ring of all formal power series and a ring of functions, which is not only a ring isomorphism but also pairs each series with a function that is one of its asymptotic sums.

Received March 1, 1956.

2. A class of linear functionals. Denote by C the collection of all functions which are asymptotic sums of formal power series (we do not allow negative exponents). If f is an asymptotic sum of the series

$$c_0 + c_1x + c_2x^2 + \dots,$$

let $L_nf = c_n$. Then L_n is a linear functional on C . If m is the first nonnegative integer such that $L_m f$ is not zero, define $\phi(f) = e^{-m}$; if $L_m f = 0$ for all m , define $\phi(f) = 0$. Then $d(f, g) = \phi(f - g)$ is a pseudo-metric on C ; two functions f and g are *asymptotically equal* (as $x \rightarrow 0+$) if $d(f, g) = 0$. These definitions are specializations to C of those given by Popken in (2).

3. Continuous linear functionals. It is natural to consider linear functionals on C which are continuous in the sense that $Lf_n \rightarrow Lf$ whenever $f_n \rightarrow f$, i.e. whenever $d(f_n, f) \rightarrow 0$. Such a functional L is constant on the classes of asymptotically equal functions.

LEMMA 1. Every continuous linear functional on C is a (finite) linear combination of functionals of the type described in §2.

Proof. Let L be a continuous linear functional, and let $L(x^m) = a_m$. If $a_m \neq 0$, let $c_m = 1/a_m$; otherwise let $c_m = 0$. Let

$$f_n(x) = c_0 + c_1x + \dots + c_nx^n.$$

Let f be any asymptotic sum of the formal series $c_0 + c_1x + c_2x^2 + \dots$ (f exists, as remarked in §1). Then $d(f_n, f) \rightarrow 0$, so by continuity $L(f_n) \rightarrow L(f)$. Since $L(f_{n+1}) = L(f_n) + 1$ whenever $a_{n+1} \neq 0$, all but a finite number of the a_m must equal zero.

Now let $d_0 + d_1x + d_2x^2 + \dots$ be any formal series, and let g be any one of its asymptotic sums. Let

$$g_n(x) = d_0 + d_1x + \dots + d_nx^n.$$

Then $d(g_n, g) \rightarrow 0$, so $L(g_n) \rightarrow L(g)$, and since all a_m are zero for $m > N$ (say), we have

$$L(g_n) = d_0a_0 + d_1a_1 + \dots + d_na_n, \quad n > N.$$

Therefore $L(g) = L(g_N)$. Recalling the definition of L_n , we prove that

$$L = a_0L_0 + a_1L_1 + \dots + a_NL_N,$$

completing the proof of the Lemma.

Conversely, any linear combination of the L_n is a continuous linear functional on C . Since this class of linear functionals has such a transparent structure, it seems to be of limited interest.

4. Asymptotic continuity.

DEFINITION. A linear functional L is *asymptotically continuous* on a subspace of C if, whenever h is in this subspace, and $h(x) = o(x^n)$ as $x \rightarrow 0+$,

for some nonnegative integer n , it follows that $h^*(t) = o(t^n)$ as $t \rightarrow 0+$, where h^* is defined by $h^*(t) = L(h(xt))$. Here L acts on functions of the positive variable x , t being a positive parameter. The transform h^* is not required to be an element of C .

It is easily verified that the functionals L_n of §2 are asymptotically continuous on C ; it follows from Lemma 1 that *any continuous linear functional on C is also asymptotically continuous*. We do not know if there exist any asymptotically continuous linear functionals on C that are not also continuous; perhaps the definitions are equivalent for functionals defined and linear on all of C . On certain subspaces the two definitions are not equivalent, and we now consider one of these subspaces. Certain other subspaces can be treated by essentially identical methods, but we give the details for only one of them here.

Let H denote the smallest subspace of C containing the functions $\{x^m\}$ ($m = 0, 1, 2, \dots$), and $\{e^{-tx}\}$ for all nonnegative real numbers t . A general element of H is simply a finite linear combination of such functions. It is easily seen that the following properties are equivalent in H ; a function possessing any one of them possesses them all:

- (1) $f(x) = o(x^n)$ as $x \rightarrow 0+$,
- (2) $f(xt) = o(x^n)$ as $x \rightarrow 0+$, for all fixed positive t ,
- (3) for some positive t , $f(xt) = o(x^n)$ as $x \rightarrow 0+$,
- (4) $L_n f = 0$ ($m = 0, 1, \dots, n$), where L_n is as defined in §2.

LEMMA 2. *If L is an asymptotically continuous linear functional on H and if $L(x^m) = (-1)^m c_m m!$ for all nonnegative integers m , then $L(e^{-tx})$ is an asymptotic sum of the formal series $c_0 + c_1 t + c_2 t^2 + \dots$*

Proof. For each n , and each t ,

$$e^{-xt} = \sum_{m=0}^n \frac{(-1)^m t^m x^m}{m!} = o(x^n), \quad x \rightarrow 0+;$$

thus by definition of asymptotic continuity (using the linearity of L) we have

$$L(e^{-xt}) - c_0 - c_1 t - \dots - c_n t^n = o(t^n), \quad t \rightarrow 0+.$$

The following lemma and the theorem in §5 both show the existence of a large class of such linear functionals.

LEMMA 3. *If $a(x)$ is a function of bounded variation possessing all moments, i.e.,*

$$\int_0^\infty x^n da(x) < \infty, \quad n = 0, 1, 2, \dots,$$

then the linear functional

$$L(f) = \int_0^\infty f(x) da(x)$$

is asymptotically continuous on H .

Proof. If $h(x) = o(x^n)$, by Taylor's theorem

$$h(x) = h^{(n+1)}(\theta x)x^{n+1}/(n+1)!$$

where $0 < \theta < 1$, θ depending on x . Since $h^{(n+1)}(x)$ is a linear combination of terms, each of which is a power of x or a function of the form e^{-cx} for non-negative c , each term of $h^{(n+1)}(\theta xt)$, for $0 < \theta < 1$, $0 < t < 1$, is dominated in magnitude either by the corresponding term of $h^{(n+1)}(x)$ or by a constant (for $x > 1$). The hypothesis that $a(x)$ possesses all moments then ensures that the integral

$$\int_0^\infty h^{(n+1)}(\theta xt) da(x)$$

is finite and bounded for $0 < t < 1$.

We then have

$$\begin{aligned} \frac{h^*(t)}{t^n} &= \frac{L(h(xt))}{t^n} = \int_0^\infty \frac{h(xt)}{t^n} da(x) \\ &= \frac{t}{(n+1)!} \int_0^\infty h^{(n+1)}(\theta xt) da(x) \rightarrow 0, \quad t \rightarrow 0+, \end{aligned}$$

proving that $h^*(t) = o(t^n)$, $t \rightarrow 0+$.

In §1, one method of constructing asymptotic sums was described. We now describe another method, based on the theory of the moment problem. This generalizes one discussed by E. Borel in (1). Given any sequence a_0, a_1, a_2, \dots of real numbers, there exists (3, p. 139) an L of the form of Lemma 3 such that $L(x^m) = a_m$. Given a formal power series

$$c_0 + c_1 t + c_2 t^2 + \dots$$

we take $a_m = (-1)^m c_m m!$; by Lemma 3 this L is asymptotically continuous on H . Since $L(x^m) = (-1)^m c_m m!$ the function $L(e^{-tx})$ is an asymptotic sum of the given series, by Lemma 2.

One of the simplest examples is provided by the convergent series

$$1 - t + t^2/2! - t^3/3! + \dots;$$

here $a_m = 1$ for all m , and we obtain $L(x^m) = 1$ for all m by taking $a(x)$ in Lemma 3 to be constant except for a unit jump at $x = 1$. Then $L(e^{-tx}) = e^{-t}$; this is actually the sum of the series. In less trivial cases, the determination of $a(x)$ may be a formidable task. Thus this method is difficult to apply; in a later paper we shall give more convenient methods based on the observation that any formal power series is the series expansion of a Schwartz distribution.

Despite these difficulties, the existence of this method shows that any linear functional L_1 on the space of all polynomials can be extended to an asymptotically continuous linear functional on H . The extension is not unique. Moreover, the following remark shows that the extension may not even be possible, if we demand continuity in the sense of §3, and this is one justification for introducing the definition of asymptotic continuity.

There exist asymptotically continuous linear functionals on H that are not continuous. For if all asymptotically continuous L were continuous, we would have

$$L(e^{-x^2}) = \sum_{n=0}^{\infty} \frac{(-1)^n L(x^n) t^n}{n!} = \sum_{n=0}^{\infty} c_n t^n,$$

the left-hand side being a function of t , and the sum being necessarily convergent. The theory of the moment problem implies that, by proper choice of such L , we can obtain any given formal power series on the right side, an obvious contradiction.

5. The main theorem. We prove a theorem which characterizes all asymptotically continuous linear functionals on H . Since L is determined completely by its values on the basis functions, let L_1 be the function defined for all nonnegative integers m by $L_1(m) = L(x^m)$, and let L_2 be a function on the positive reals defined by $L_2(t) = L(e^{-t^2})$. Then L can be identified with the pair $\{L_1, L_2\}$, since any arbitrary pair of such functions determines uniquely a linear functional L (not necessarily asymptotically continuous). Since $L_2(t)$ may be discontinuous, it is an easy corollary of this theorem that not every asymptotically continuous linear functional on H is of the type described in Lemma 3.

THEOREM. $L = \{L_1, L_2\}$ is asymptotically continuous on H if and only if $L_2(t)$ is an asymptotic sum of the formal power series

$$\sum_{m=0}^{\infty} \frac{(-1)^m L_1(m) t^m}{m!}.$$

Proof. The necessity is a rewording of Lemma 2, so we only need prove the sufficiency. Any element of H may be written in the form

$$h(x) = \sum_{n=0}^p a_n x^n + \sum_{m=1}^r b_m e^{-t_m x^2},$$

and by the remark preceding Lemma 2, if $h(x) = o(x^n)$ as $x \rightarrow 0+$, we compute

$$a_m = - \sum_{q=1}^r \frac{(-1)^m t_q^m b_q}{m!}, \quad m = 0, 1, \dots, n.$$

Thus we have

$$L(h(xt)) = \sum_{q=1}^r b_q \left[L_2(t_q t) - \sum_{m=0}^n \frac{(-1)^m t_q^m L_1(m) t^m}{m!} \right] + \sum_{m=n+1}^p a_m L_1(m) t^m.$$

If $L_2(t_q t)$ is an asymptotic sum of

$$\sum_{m=0}^{\infty} \frac{(-1)^m L_1(m) (t_q t)^m}{m!},$$

each term in the equation is $o(t^n)$, $t \rightarrow 0+$, proving that L is asymptotically continuous.

REFERENCES

1. E. Borel, *Leçons sur les séries divergentes* (Paris, 1928).
2. J. Popken, *Asymptotic expansions from an algebraic standpoint*, Nederl. Akad. Wetensch. Proc. Ser. A 56 = Indagationes Math. 15 (1953), 131-143.
3. D. W. Widder, *The Laplace Transform* (Princeton, 1946).

University of British Columbia

ON WARD'S PERRON-STIELTJES INTEGRAL

RALPH HENSTOCK

Introduction. In the paper (5), Ward defines an integral of Perron type of a finite function f with respect to another finite function g , where g need not be of bounded variation. There arise two problems, (a) and (b) below, that have not been dealt with in (5).

If $f = j$ at a countable number of points everywhere dense in (a, b) , where f and j are both integrable with respect to g , then $f - j$ can be nonzero on a large set of points of (a, b) . For example, if g is continuous and of bounded variation the countable number of points can be neglected in the integration and we can have $f \neq j$ everywhere else. But g is more rigidly fixed when we know its values on an everywhere dense set, if the integral exists. For example, if g is of bounded variation, and so continuous except at an at most countable set of points, we can only vary the values of g at a countable set of points. More generally, we have problem

(a) *If f is integrable with respect to g , and with respect to h , over the closed interval $[a, b]$, where $g = h$ at points everywhere dense in $[a, b]$, what are the properties of the difference $g - h$ and the set of points where the difference is not zero?*

This question is partially answered by Theorems 1 and 2, and we obtain the following result.

Let \bar{E}_ϵ be the closure of the set of u for which

$$(1) \quad |g(u) - h(u)| \geq \epsilon, \quad a \leq u \leq b.$$

Then f must be VBG and continuous on¹ \bar{E}_ϵ , and $mf(\bar{E}_\epsilon) = 0$.

However, if f is integrable with respect to g in $[a, b]$, and if $g - h$ satisfies (i) and is 0 at an everywhere dense set of points in $[a, b]$, it does not follow that f is integrable with respect to h in $[a, b]$. For example, take $g = 0$ and suppose that each set E_ϵ contains only a finite number of points and so has no limit-points. Then every function f is trivially VBG and continuous on $\bar{E}_\epsilon = E_\epsilon$, and $f(\bar{E}_\epsilon)$ contains only a finite number of points. But if the set of points where $h \neq 0$ does not satisfy Theorem 3 (9), (10), (11), with j replaced by h , it follows by Theorem 3 that there is a finite function f for which the Perron-Stieltjes integral of f with respect to h over $[a, b]$ does not exist. See the example of Theorem 5 (38) in §4.

There is another question of integrability, namely,

(b) *What are the properties of g in order that all bounded Baire² functions f are integrable with respect to g in $[a, b]$?*

Received October 6, 1955.

¹I.e., when we use only the points of \bar{E}_ϵ .

²A Baire (Borel-measurable) function is any function that can be obtained from continuous functions by using repeated limits.

Question (b) is partially answered in (2), Theorem 2, and we give the complete answer in Theorem 3 of the present paper.

1. Notation. We suppose that all functions considered are defined and finite in $a < u < b$, this interval being denoted by $[a, b]$. The existence of an integral or limit is taken to mean its existence as a finite number. If the limits exist,

$$f(u-) = \lim_{v \rightarrow u, a < v < u < b} f(v), \quad f(u+) = \lim_{v \rightarrow u, a < u < v < b} f(v).$$

Integral signs preceded by (LS), (PS), denote respectively the Lebesgue-Stieltjes and Perron-Stieltjes integrals, and we put

$$P(v, w) \equiv P(f, g; v, w) \equiv (PS) \int_v^w f(u) dg(u),$$

$f(E) \equiv \{f(u) : u \in E\}$ where E is a set contained in $[a, b]$. A point v in $[a, b]$ is a point of *infinite variation* on $[a, b]$ of the function f if, for each open interval (ξ, η) containing v , the function f is not of bounded variation on

$$[\xi, \eta] \cap [a, b].$$

It follows that the set W of points of infinite variation on $[a, b]$ of f is closed. For if v is not in W there is an open interval (ξ, η) containing v , such that f is of bounded variation on

$$[\xi, \eta] \cap [a, b],$$

and then (ξ, η) is contained in CW .

The symbols E' , \bar{E} , CE , mE denote respectively the derived set, the closure, the complement, and the measure of a set E in $[a, b]$. The interior of E is the largest open set contained in E .

2. The examination of question (a)

THEOREM 1. If $P(f, g; a, b)$ and $P(f, h; a, b)$ exist, and if $g = h$ at points everywhere dense in $[a, b]$, then for all v, w in $a < v < w < b$,

$$P(f, g; v, w) = P(f, h; v, w) + [f(g - h)]_v^w.$$

Proof. It is enough to assume that $h \equiv 0$, so that $g = 0$ at points everywhere dense in $[a, b]$. Let M_1 and M_2 be a major and a minor function, in Ward's sense, of f with respect to g in $[a, b]$ and take u in $[a, b]$. Then there is a $\delta_1(u) > 0$ depending on u , M_1 , M_2 , such that

$$(2) \quad [M_1]_u^\xi > f(u)[g]_u^\xi > [M_2]_u^\xi, \quad 0 < \xi - u < \delta_1(u),$$

$$(3) \quad [M_1]_u^\xi < f(u)[g]_u^\xi < [M_2]_u^\xi, \quad 0 > \xi - u > -\delta_1(u).$$

As in (2), §2, the proof of Theorem 1, we can prove that in each $[v, w]$ there is a finite number of points

$v = \alpha_0 = u_1 < \alpha_1 < \dots < \alpha_n = u_n = w$, $\alpha_{p-1} < u_p < \alpha_p$ ($p = 2, \dots, n-1$), such that

$$g(\alpha_p) = 0 \quad (p = 1, \dots, n-1), \quad \alpha_p - \alpha_{p-1} < \delta_1(u_p) \quad (p = 1, \dots, n).$$

Thus (2), (3) are satisfied with $u = u_p$, $\xi = \alpha_p$, and $u = u_p$, $\xi = \alpha_{p-1}$, respectively, and we obtain

$$[M_1]_v^w = \sum_{p=1}^n [M_1]_p > [fg]_v^w > \sum_{p=1}^n [M_2]_p = [M_2]_v^w,$$

where $[M]_p$ stands for

$$M(\alpha_p) - M(\alpha_{p-1}).$$

Thus as $P(v, w)$ exists, the Theorem must be true for $h \equiv 0$, and so generally.

THEOREM 2. If, for all u in $a < u < b$,

$$(4) \quad P(f, g; a, u) = [fg]_a^u,$$

then (5) f is VBG and continuous on \bar{E}_ϵ , and (6) $mf(\bar{E}_\epsilon) = 0$, where E_ϵ is the set of u for which

$$|g(u)| > \epsilon, \quad a < u < b, \quad \epsilon > 0.$$

COROLLARY. If (4) is true, and if \bar{E}_ϵ contains an interval $[\xi, \eta]$ for some $\epsilon > 0$, then f is constant in $[\xi, \eta]$.

From Theorem 2 Corollary we can easily prove Theorem 1 of (2).

To prove Theorem 2 let $a < u < v < b$ and let M_3, M_4 be arbitrary major and minor functions of f with respect to g in Ward's sense, and write $\chi_1 \equiv M_3 - M_4$. Then χ_1 is monotone increasing. Now, for fixed u and for sufficiently small and positive $v - u$, both functions

$$f(u)[g]_u^v, \quad P(u, v)$$

lie between

$$[M_3]_u^v, \quad [M_4]_u^v,$$

so that

$$|P(u, v) - f(u)[g]_u^v| < [\chi_1]_u^v.$$

Substituting in the value of $P(u, v)$ from (4) we obtain

$$|g(v)[f]_u^v| < [\chi_1]_u^v.$$

Hence there is a $\delta_2(u) > 0$ such that if

$$u \in E_\epsilon', v \in E_\epsilon, \quad 0 < v - u < \delta_2(u),$$

we have

$$(7) \quad |[f]_u^v| < \epsilon^{-1}[\chi_1]_u^v.$$

Similarly for $v < u$. If

$$w \in \bar{E}_\epsilon, \quad 0 < w - u < \delta_2(u),$$

then there is a v satisfying

$$v \in E_\epsilon, \quad 0 < v - u < \delta_2(u),$$

and arbitrarily near to w , so that by (7),

$$\begin{aligned} |[f]_w^w| &\leq |[f]_u^u| + |[f]_v^v| \leq \epsilon^{-1}[\chi_1]_u^\epsilon + \epsilon^{-1}[\chi_1]_v^\epsilon, \\ (8) \quad |[f]_w^w| &\leq \epsilon^{-1}[\chi_1]_u^{\epsilon+} \leq \epsilon^{-1}[\chi_1]_u^\epsilon, \\ \limsup_{w \rightarrow u} |[f]_w^w| &\leq \epsilon^{-1}[\chi_1]_u^\epsilon, \quad \lim_{w \rightarrow u} f(w) = f(u), \end{aligned}$$

as $\chi_1(b) - \chi_1(a)$ is arbitrarily small.

Similar results hold for

$$w < u, u \in E'_\epsilon, w \in \bar{E}_\epsilon, w \rightarrow u,$$

so that f is continuous when we only use the points of the derived set of E_ϵ . As the other points of \bar{E}_ϵ are isolated, f is continuous on \bar{E}_ϵ .

To show that f is VBG on \bar{E}_ϵ we use the method of the first part of the proof of (5, p. 592, Lemma 6) and we employ only points of \bar{E}_ϵ . The relevant inequality is the first one in (8).

To prove (6) we first add $\theta(u - a)$ to $\chi_1(u)$ if necessary, to ensure that χ_1 is strictly increasing. The constant $\theta > 0$ can be arbitrarily small. Then as in (5, p. 581, Lemma 3) we prove from the first inequality of (8), and the similar inequality when $w < u$, that

$$m^*f(\bar{E}_\epsilon) \leq 2\epsilon^{-1}[\chi_1]_a^\epsilon,$$

where m^* denotes outer measure. The factor 2 occurs because of the $w+$ in (8). As the right-hand side is arbitrarily small we obtain (6).

To prove the Corollary we note that by (5), f is continuous on $[\xi, \eta]$. Thus if $f([\xi, \eta])$ contains two distinct points it contains the whole interval between the points. This is impossible by (6).

3. The integrability of Perron-Stieltjes integrals. In this section we prove two theorems, completely answering question (b). We begin with a lemma needed in the proof of the converse of Theorem 3.

LEMMA. Let F be a sequence $\{I_n\}$ of open intervals, and let H_p be the set of points of $[a, b]$ lying in at most p intervals of F . Then all the intervals I_n covering the points of H_p can be put into at most $3p$ sets of non-overlapping intervals.

We can define a sequence $\{\xi_\epsilon\}$ of points of H_p such that their closure contains H_p . Each interval I_n covering a point of H_p will then also cover at least one ξ_ϵ , and conversely. Thus we need only consider the intervals covering the ξ_ϵ .

We put the q th interval of the sequence $\{I_n\}$ that covers ξ_1 into the set S_q . Then $1 < q < p$, as ξ_1 lies in H_p . Suppose that the intervals I_n covering ξ_1, \dots, ξ_{r-1} have been arranged into sets $S_q (1 < q < 3p)$ of non-overlapping intervals, and let ξ_r lie between ξ_s and ξ_t for $s < r, t < r$, with no $\xi_q (q < r)$ between ξ_s and ξ_t . Then there are at most p intervals I_n covering ξ_s , and at most p intervals I_n covering ξ_t , so that at least p of the sets S_1, \dots, S_{3p} , say T_1, \dots, T_p , will be free from intervals I_n that cover ξ_s or ξ_t , and so will contain no interval lying in (ξ_s, ξ_t) . The intervals I_n covering ξ_r that have not already been put into sets S_q , cannot cover ξ_s nor ξ_t , and so must lie between ξ_s and ξ_t . We can therefore put these intervals into some or all of the sets T_1, \dots, T_p .

Similarly if

$$\xi_r < \min_{q < r} \xi_q \text{ or } \xi_r > \max_{q < r} \xi_q,$$

in which case one of ξ_s, ξ_t is missing. Hence the result is true for ξ_1, \dots, ξ_r . It is true for ξ_1 and hence true in general.

THEOREM 3. *If, for a given function j , for all bounded Baire functions f defined in $[a, b]$, and for all u in $[a, b]$, the integral $P(f, j; a, u)$ exists equal to*

$$[fj]_a^u,$$

then the set of points u in $a < u < b$, where $j(u) \neq 0$, can be divided into two sequences $\{u_n\}$ and $\{d_n\}$, with the properties

$$(9) \quad \sum_{n=1}^{\infty} |j(u_n)| < \infty;$$

(10) *surrounding each d_n there is an open interval $I(d_n) \equiv (\underline{d}_n, \bar{d}_n)$ contained in (a, b) such that each point of $[a, b]$ can lie in at most finite number of the $I(d_n)$;*

(11) *there is a monotone increasing bounded function χ such that*

$$\chi(\bar{d}_n+) - \chi(d_n) > |j(d_n)|, \quad \chi(d_n) - \chi(\underline{d}_n-) > |j(d_n)|.$$

Conversely, if j satisfies (9), (10), (11), and if f is bounded in $[a, b]$, then $P(f, j; a, u)$ exists and is equal to

$$[fj]_a^u,$$

for all u in $a < u < b$.

To begin the proof of the first part of Theorem 3 we replace g by j in Theorem 2, obtaining from (5) that f is continuous on \bar{E}_ϵ , where E_ϵ is the set in which $|j| > \epsilon$. But, for each u in $[a, b]$, the set of bounded Baire functions f includes the function equal to 0 in $[a, u]$, equal to 1 at u , and equal to 2 in $(u, b]$. Hence each point of \bar{E}_ϵ must be isolated, and E_ϵ is finite. This is true for each $\epsilon > 0$. Hence taking $\epsilon^{-1} = 1, 2, \dots$, we obtain

(12) $j \neq 0$ only at a countable set of points $\{w_n\}$,

(13) $j(w_n) \rightarrow 0$ as $n \rightarrow \infty$.

Also, as E_1 is finite,

(14) j is bounded.

We now wish to find a strictly increasing function χ and a function $\delta > 0$ defined for all u in $a < u < b$, such that for $u - \delta < w < u < v < u + \delta$, $a < w < v < b$,

$$(15) \quad [\chi]_u^v > |j(v)|,$$

$$(16) \quad [\chi]_w^u > |j(w)|.$$

There is in Ward's sense a major function $P(f, j; a, u) + \chi_2(u)$ of f with respect to j in $[a, b]$, where χ_2 is monotone increasing and bounded in $[a, b]$, with $\chi_2(a) = 0$. Thus, if we substitute in the value of $P(f, j; a, u)$, we find that for $a < u < b$ and for some $\delta_3 = \delta_3(u) > 0$, using Ward's definition of a major function,

$$(17) \quad [\chi_2]_u^v > j(v)[f]_u^v \quad (u < v < u + \delta_3, a < v < b),$$

$$(18) \quad [\chi_2]_w^u > j(w)[f]_w^u \quad (u > w > u - \delta_3, a < w < b).$$

We now take $f = -\operatorname{sgn} j$, where $\operatorname{sgn} a \equiv |a|/a$ ($a \neq 0$), $\operatorname{sgn} 0 = 0$. Then if χ_3, δ_4 are the corresponding χ_2, δ_3 , and if the u of (17) does not lie in $\{w_n\}$, so that $j(u) = 0, f(u) = 0$, we obtain, for $u < v < u + \delta_4, a < v < b$,

$$(19) \quad [\chi_3]_u^v > |j(v)|.$$

Similarly let χ_4, δ_5 be the corresponding χ_2, δ_3 when for f we take $\operatorname{sgn} j$, and let the u of (18) lie outside the sequence $\{w_n\}$ so that $j(u) = 0, f(u) = 0$. Then

$$(20) \quad [\chi_4]_w^u > |j(w)|, \quad u > w > u - \delta_5, a < w < b.$$

By (13), $j(w_n \pm) = 0$. Thus if we put

$$\chi_5(u) = \sum_{w_n < u} 2^{-n} (u \notin \{w_n\}) = \chi_3(w_p-) + 2^{-2p} \quad (u = w_p, p = 1, 2, \dots)$$

we obtain

$$\chi_5(w_p+) - \chi_5(w_p) = 2^{-2p} > 0 = |j(w_p+)|,$$

$$\chi_5(w_p) - \chi_5(w_p-) = 2^{-2p} > 0 = |j(w_p-)|,$$

and there is a number $\delta_p = \delta(w_p)$ such that $\chi_5(u)$ satisfies (15) and (16) at $u = w_p$, with χ replaced by χ_5 and δ by δ_p .

Using (19), (20) also, we see that to obtain (15), (16) for all u in $a < u < b$ and a strictly increasing function χ , we need only take

$$\chi(u) = \chi_3(u) + \chi_4(u) + \chi_5(u) + u - a.$$

We now define the points d_n in (a, b) as those for which

$$(21) \quad |j(d_n)| > \chi(d_n+) - \chi(d_n), \quad |j(d_n)| > \chi(d_n) - \chi(d_n-).$$

The other points $\{u_n\}$ of $\{w_n\}$ then give

$$\sum_{n=1}^m |j(u_n)| \leq \sum_{n=1}^m \{ \chi(u_n+) - \chi(u_n-) \} \leq [\chi]_a^b < \infty,$$

so that (9) is satisfied.

If $u < d_n < u + \delta(u)$ for some u, d_n , we have (15) with $v = d_n$. Let q_n be the upper bound of all $u < d_n$ satisfying (15) for fixed $v = d_n$. If there is no such u , put $q_n = a$. Then

$$(22) \quad \chi(d_n) - \chi(q_n-) \geq |j(d_n)|,$$

while if $d_n > u > q_n$, we have

$$(23) \quad \chi(d_n) - \chi(u) < |j(d_n)|.$$

By (14), j is bounded, so that we can take a convenient finite value for $\chi(a-)$ to fit the cases when $q_n = a$. From (21), (22), $q_n < d_n$.

Similarly we can define $\bar{d}_n > d_n$ such that

$$(24) \quad \chi(\bar{d}_n+) - \chi(d_n) \geq |j(d_n)|,$$

while if $d_n < u < \bar{d}_n$, we have

$$(25) \quad \chi(u) - \chi(d_n) < |j(d_n)|.$$

Results (22), (24) prove (11). We now suppose that (10) is false, so that a point u of $[a, b]$ lies in an infinity of the open intervals

$$I(d_n) \equiv (q_n, \bar{d}_n) \subseteq (a, b).$$

Obviously $u \neq a, u \neq b$. Also by (23), (25), (13),

$$\chi(\bar{d}_n-) - \chi(q_n+) \leq 2|j(d_n)| \rightarrow 0$$

as $n \rightarrow \infty$. Hence as χ is strictly increasing, $q_n \rightarrow u$ and $\bar{d}_n \rightarrow u$, for the subsequence of n for which $q_n < u < \bar{d}_n$. Hence the corresponding subsequence of $\{d_n\}$ also tends to u , so that for certain $v \rightarrow u$,

$$|\chi(v) - \chi(u)| < |j(v)|.$$

This result contradicts (15) or (16). Hence (10) is true, and the first part of Theorem 3 has been proved.

We now prove the converse. Let the discontinuities of χ in $[a, b]$ occur at the points $v_n (n = 1, 2, \dots)$. Then we have

$$\sum_{n=1}^m \{ \chi(v_n+) - \chi(v_n-) \} \leq [\chi]_a^b < \infty,$$

so that, given $\epsilon > 0$, there is an integer n_0 such that

$$(26) \quad \sum_{n=n_0}^m \{ \chi(v_n+) - \chi(v_n-) \} < \epsilon.$$

Then there is an integer n_1 such that, for $n > n_1$, d_n is not one of the points $v_q (q = 1, \dots, n_0 - 1)$.

We now let F in the Lemma be the family of intervals $I(d_n)$, and we take p so large that

$$(27) \quad m \chi\{[a, b] - H_p\} < \epsilon.$$

This is possible since by (10),

$$[a, b] = \bigcup_{n \geq 0} H_p.$$

By the Lemma there are $3p$ sets S_q of non-overlapping intervals $I(d_n)$ that together cover $H_p - H_0$. There is an integer $t > n_1$, and depending on ϵ , such that for each q in $1 \leq q \leq 3p$,

$$(28) \quad \sum \{ \chi(\bar{d}_n+) - \chi(d_n-) \} < \epsilon/(3p),$$

where the sum is taken over those intervals of S_q with $n > t$, as the sum for $n > 0$ is not greater than $\chi(b) - \chi(a)$. The integer t can also be chosen, by (9), so that

$$(29) \quad \sum_{n > t} |j(u_n)| < \epsilon.$$

Let S be the set formed from those intervals of the S_q with $n > t$ and $1 \leq q \leq 3p$. Then

$$\{[a, b] - H_p\} \cup S$$

is a union of intervals. For if u lies in $[a, b] - H_p$ let J be the intersection of the first $(p+1)$ intervals $I(d_n)$ covering u . Then J is open and contains u , and

$$J \subseteq [a, b] - H_p.$$

We add an at most countable number of points, if necessary, to obtain from $\{[a, b] - H_p\} \cup S$ a union U of open non-abutting intervals, and we put

$$(30) \quad \chi_\epsilon(u) \equiv \sum_1 \{ \chi(\beta+) - \chi(\alpha-) \} + \epsilon(u-a)/(b-a) + \sum_2 2|j(u_n)|,$$

where \sum_1 denotes the summation over the intervals (α, β) of $U \cap (a, u)$, changing $\beta+$ to β if $\beta = u$; and \sum_2 denotes the summation over all $n > t$ such that $u_n < u$, adding $|j(u_n)|$ if $p > t$ and $u = u_p$. Then χ_ϵ is strictly increasing, and from (26), (27), (28), (29),

$$(31) \quad [\chi_\epsilon]_a^b < 6\epsilon.$$

Now, by definition, the points of H_0 are not covered by any interval $I(d_n)$. If $n > t$ and if $I(d_n)$ covers a point of $H_p - H_0$, then $I(d_n)$ will lie in one of the S_q , and so in S , and so in U . It follows that $\chi(d_n) - \chi(d_n-)$ will occur in \sum_1 for $u = d_n$. If $n > t$ and if $I(d_n)$ does not cover a point of $H_p - H_0$, then $I(d_n)$ will lie entirely within $[a, b] - H_p$, and so in U , and again, $\chi(d_n) - \chi(d_n-)$ will occur in \sum_1 for $u = d_n$. Thus by (30),

$$(32) \quad \chi_\epsilon(d_n) - \chi_\epsilon(d_n-) \geq \chi(d_n) - \chi(d_n-) \geq |j(d_n)| \quad (n > t).$$

Similarly for the result with \bar{d}_n+ , so that χ_ϵ satisfies (11) for all $n > t$.

Now each point u of $[a, b]$ lies in at most finite number of the $I(d_n)$, say $I(\xi_1), \dots, I(\xi_r)$, where ξ_1, \dots, ξ_r depend on u . Let the sequence $\{\eta_n\}$ include all points of the sequences $\{u_n\}$, $\{d_n\}$, $\{d_n\}$, $\{\bar{d}_n\}$, and let u be outside $\{\eta_n\}$. We take $\delta_6 = \delta_6(u) > 0$ so that $(u - \delta_6, u + \delta_6)$ does not include

$$u_1, \dots, u_t, d_1, \dots, d_t, \xi_1, \dots, \xi_r.$$

Then by (32), for $u < d_n < \min(b, u + \delta_6)$,

$$\chi_6(d_n) - \chi_6(u) > \chi(d_n) - \chi(d_n -) > |j(d_n)|,$$

since $d_n > u$. If u_n lies in $u < u_n < \min(b, u + \delta_6)$ then $n > t$, and by (30),

$$\chi_6(u_n) - \chi_6(u) > |j(u_n)|.$$

If v is neither in $\{u_n\}$ nor in $\{d_n\}$ then for $u < v < \min(b, u + \delta_6)$,

$$\chi_6(v) - \chi_6(u) > 0 = |j(v)|.$$

Hence, if u is outside $\{\eta_n\}$,

$$(33) \quad \chi_6(v) - \chi_6(u) > |j(v)|, \quad u < v < \min(b, u + \delta_6).$$

Similarly for all v in $u > v > \max(a, u - \delta_6)$. To deal with the case when $u = \eta_n$ for some n , we put

$$\begin{aligned} \chi_7(u) &= \chi_6(u) + \sum_{\eta_n < u} \epsilon 2^{-n} & (u \notin \{\eta_n\}), \\ \chi_7(\eta_p) &= \chi_7(\eta_p -) + \epsilon 2^{-2p} & (p = 1, 2, \dots). \end{aligned}$$

As in the part of the proof that follows (20), we obtain a strictly increasing function χ_7 satisfying (33) for all u , and, for suitable $\delta_7 > 0$, for

$$u < v < \min(b, u + \delta_7),$$

and similarly for $v < u$. By (31),

$$(34) \quad [\chi_7]_a^b < 7\epsilon.$$

Now suppose that $|f| \leq A$. We put

$$M_5(u) = [fj + 2A\chi_7]_a^u.$$

Then from (33),

$$\begin{aligned} [M_5]_u^v - f(u)[j]_u^v &= [f]_u^v j(v) + 2A[\chi_7]_u^v \\ &> [f]_u^v j(v) + 2A|j(v)| > 0 \quad (u < v < \min(b, u + \delta_7)). \end{aligned}$$

The inequalities are reversed when $u > v > \max(a, u - \delta_7)$, so that M_5 is a major function, in Ward's sense, for f with respect to j in $[a, b]$. Similarly

$$M_6(u) = [fj - 2A\chi_7]_a^u$$

is a minor function, and by (34),

$$M_6(b) - M_5(b) = 4A[\chi_7]_a^b < 28A\epsilon.$$

By choice of $\epsilon > 0$ this can be made arbitrarily small. Hence there exists

$$P(f, j; a, u) = [fj]_a^u$$

proving the converse in Theorem 3.

THEOREM 4. *If, for a given function g , and for all bounded Baire functions f in $[a, b]$, the integral $P(f, g; a, b)$ exists, then*

(35) $g(u-)$ exists in $a < u \leq b$, $g(u+)$ exists in $a \leq u < b$, and both are of bounded variation in those ranges; and the function j satisfies Theorem 3(9), (10), (11), where

$$(36) \quad j(a) = g(a) - g(a+), \quad j(b) = g(b) - g(b-),$$

$$j(u) = g(u) - \frac{1}{2}\{g(u+) + g(u-)\} \quad (a < u < b).$$

Conversely, if g satisfies (35), and if the j defined by (36) satisfies Theorem 3(9), (10), (11), and if f is a bounded Baire function in $[a, b]$, then $P(f, g; a, b)$ exists and is equal to

$$\{g(b) - g(b-)\}f(b) + \{g(a+) - g(a)\}f(a) + \sum_{a < u < b} f(u)\{g(u+) - g(u-)\} \\ + (LS) \int_a^b f(u) dg_e(u),$$

where

$$g_e(v) = g(v-) - \sum_{a < u < v} \{g(u+) - g(u-)\} \quad (a < v \leq b), \quad g_e(a) = g(a+).$$

The result (35) is proved in (2), Theorem 2, using only the hypotheses of the present Theorem 4. From (35) we see that $g - j$ is of bounded variation in $[a, b]$, so that $P(f, g - j; a, b)$ exists. By hypothesis $P(f, g; a, b)$ exists. Hence so does $P(f, j; a, b)$. Also, from (35),

$$\lim_{w \rightarrow u-} g(w-) = g(u-), \quad \lim_{w \rightarrow u-} g(w+) = g(u-),$$

so that from (36), $j(u-) = 0$. Similarly $j(u+) = 0$. If E_ϵ is the set in $a < u \leq b$ where $j > \epsilon > 0$, and if E_ϵ has a limit-point ξ , then

$$\limsup_{w \rightarrow \xi} j(w) > \epsilon.$$

This contradicts $j(\xi-) = 0 = j(\xi+)$, so that E_ϵ has no limit-points and so must contain only a finite number of points. Thus taking $\epsilon = n^{-1}$ ($n = 1, 2, \dots$), the set where $j > 0$ is at most countable. Similarly the set where $j < 0$ is at most countable. Hence by Theorem 1,

$$P(f, j; a, u) = [fj]_a^u$$

so that the first part of Theorem 3 completes the first part of Theorem 4.

To prove the converse in Theorem 4 we need only use the converse in Theorem 3 and the fact that $g - j$ is of bounded variation in $[a, b]$, and (4, pp. 208-209, Theorem 8.1)).

4. The points of infinite variation of j . We now suppose that

$$(37) \quad j(u-) = 0 \quad (a < u \leq b), \quad j(u+) = 0 \quad (a \leq u < b).$$

Let T_1 be the union of the interiors of all closed intervals J contained in $[a, b]$, such that $P(f, j; J)$ exists for all bounded Baire functions f , adding one or both of a, b to T_1 according as one or both of $[a, a + \epsilon]$, $[b - \epsilon, b]$ are intervals J for some $\epsilon > 0$. Also put $T = CT_1 \cap [a, b]$. Let W be the set of points of infinite variation of j .

THEOREM 5. *If J is a closed interval, there is a function j satisfying (37), such that*

$$(38) \quad J = W, J = T.$$

If Q is a closed nowhere dense set, there is a function j satisfying (37), such that

$$(39) \quad T = W = Q,$$

and there is another function j satisfying (37), such that

$$(40) \quad T = \phi, W = Q,$$

where ϕ is the empty set.

We begin by supposing that

(41) *the set of points $\{v_n\}$ in $[a, b]$ can be put into one-one correspondence with the points $(2q+1)2^{-p}$ ($0 \leq q < 2^{p-1}$; $p = 1, 2, \dots$), the order of the points being preserved.*

Then we define $j(v_n) = p^{-1}$ when v_n corresponds to $(2q+1)2^{-p}$, and $j(u) = 0$ when u is outside $\{v_n\}$. Such a j satisfies (37), as only a finite number of $j(v_n)$ are greater than any given positive ϵ . If a χ exists satisfying Theorem 3(10), (11), we can suppose that

$$(42) \quad [\chi]_a^b = B, \quad [\chi]_u^v \geq v - u,$$

for all $a \leq u < v \leq b$. Then the set of intervals $I(d_n)$ for which

$$\chi(\bar{d}_n+) - \chi(\bar{d}_n-) \geq 2/p$$

must be such that any non-overlapping and non-abutting subset has at most $\frac{1}{2}pB$ members. Hence any non-overlapping subset has at most pB members. The points of $\{v_n\}$ that are not in $\{d_n\}$ are points $\{u_n\}$ satisfying Theorem 3(9). It follows that for some integer r , there is a point d_{01} in $\{d_n\}$ with

$$\chi(\bar{d}_{01}+) - \chi(\bar{d}_{01}-) \geq 2/r$$

such that $I(d_{01})$ contains at least two different points ξ_1, ξ_2 of $\{v_n\}$ corresponding to points $(2q+1)2^{-r}$ with the given r . Hence

$$Q_1 \equiv I(d_{01}) \cap \{v_n\} \cap (\xi_1, \xi_2)$$

is not empty, as there are points of $\{v_n\}$ between each two points of $\{v_n\}$ by (41). Since ξ_1, ξ_2 lie at a positive distance from the ends of $I(d_{01})$, and since

$$\bar{d}_n - \underline{d}_n \leq \chi(\bar{d}_n +) - \chi(\underline{d}_n -) \rightarrow 0$$

as $n \rightarrow \infty$, by (42), (10), and the bounded variation of χ , there is an n_2 such that if $n > n_2$ and $d_n \in Q_1$ then

$$I(d_n)' \subseteq I(d_{01}).$$

We can now repeat the construction, defining d_{02}, d_{03}, \dots , and

$$I(d_{01}) \supseteq I(d_{02}) \supseteq \dots \supseteq I(d_{0n}) \supseteq \dots$$

As $\{d_{0n}\}$ is a subsequence of $\{d_n\}$ we have $\bar{d}_{0n} - \underline{d}_{0n} \rightarrow 0$ as $n \rightarrow \infty$, and hence for a point u in (a, b) , $I(d_{0n}) \rightarrow u$. This u lies in an infinity of the intervals $I(d_n)$, contrary to (10). Hence in this case there is no χ satisfying Theorem 3(10), (11), so that for some bounded Baire function f , $P(f, j; a, b)$ cannot exist.

A similar result is true for each interval J containing points of $\{v_n\}$ in its interior, by (41). Hence

$$(43) \quad T \supseteq \{v_n\}',$$

since by (41) each point of $\{v_n\}'$ is the limit-point of a sequence of intervals of T .

To prove (38) let J be the interval $[\alpha, \beta]$. Then the points

$$v_n = \alpha + (\beta - \alpha)(2q + 1)2^{-p} \quad (0 \leq q \leq 2^{p-1}; p = 1, 2, \dots)$$

will satisfy (41), and by (43),

$$\{v_n\}' = J = T.$$

To prove (39) we take the points v_n to be the centres of the intervals I_n complementary to Q in $[a, b]$. That $\{v_n\}$ so defined satisfies (41), can be shown by {3, p. 57, Proposition 20}. Then by (43),

$$T = \{v_n\}' = Q,$$

and (39) is proved.

To prove (40) let d_{1n} be the centre of the n th interval $J_n \equiv (\alpha_n, \beta_n)$ complementary to Q in $[a, b]$. Next, let d_{2n1} and d_{2n2} be the centres of (α_n, d_{1n}) and (d_{1n}, β_n) , respectively, calling these two points the *points of the second stage*. We continue this process of continued bisection to the stage n^2 . If d_{pnq} is a point of the p th stage in J_n put $j(d_{pnq}) = n^{-2}2^{-p}$, with $(\underline{d}_{pnq}, \bar{d}_{pnq})$ as the $(p-1)$ th stage interval with centre d_{pnq} . If this is done for $1 \leq p \leq n^2$ ($n = 1, 2, \dots$) with $j = 0$ elsewhere, and if

$$\chi(\bar{d}_{pnq}) - \chi(\underline{d}_{pnq}) = n^{-2}2^{-p}$$

we have

$$\chi(\beta_n) - \chi(\alpha_n) = n^{-2}/2,$$

and the construction of a strictly increasing χ satisfying the required conditions is possible. Each point of $[a, b]$ lies in an at most finite number of the $I(d_{pnq})$, as it lies in at most n^2 in the interval J_n . Finally, over all the points d_{pnq} in J_n ,

$$\sum |j(d_{pnq})| = \frac{1}{2}.$$

Thus T is empty and $W = Q$, proving (40).

THEOREM 6. *Let j satisfy (37), with T, W as defined just before Theorem 5. Then:*

(44) T is perfect;

(45) $W \supseteq T$;

(46) The interior of W is contained in T ;

(47) If $Q \subseteq R$ are two perfect sets in $[a, b]$ with the same interior, there is a j such that $T = Q, W = R$;

(48) In order that T should be empty, it is necessary but not sufficient that the set of points $\{d_n\}$ of Theorem 3 should be scattered.³

COROLLARY 1. *If W is at most countable then T is empty and $P(f, j; a, b)$ exists.*

COROLLARY 2. *No structural property of W can be both necessary and sufficient for T to be empty.*

By construction, T is closed. Thus to prove (44) we have only to show that T has no isolated points. Suppose on the contrary that v is an isolated point of T . Then there are points α, β , such that $\alpha < v < \beta$, with $[\alpha, v)$ and $(v, \beta]$ in T_1 . Putting

$$v_n = v - (v - \alpha)/(n + 1),$$

we see that

$$P_n = P(f, j; v_n, v_{n+1})$$

exists for each n and each bounded Baire function f . By hypothesis $j = 0$ except at an at most countable set of points, so that by Theorem 1,

$$P_n = f(v_{n+1})j(v_{n+1}) - f(v_n)j(v_n).$$

Hence for each $\epsilon > 0$ there is an increasing function χ_ϵ such that

$$[fj]_n^u + \chi_\epsilon(u), [fj]_n^u - \chi_\epsilon(u)$$

are a major and a minor function, respectively, in $\alpha \leq u < v$, in Ward's sense, with

$$\chi_\epsilon(v_{n+1}) - \chi_\epsilon(v_n) \leq \epsilon 2^{-n}, \quad \chi_\epsilon(u) - \chi_\epsilon(\alpha) \leq 2\epsilon.$$

If we set $\chi_\epsilon(v) - \chi_\epsilon(v-) = \epsilon$, then as f is bounded, say by A , and $j(v-) = 0$, we have

$$[\chi_\epsilon]_n^v > \epsilon > 2A|j(u)| > [fj]_n^u j(u)$$

³"Zerstreute" (F. Hausdorff), "separierte" (G. Cantor), "clairsemé" (A. Denjoy).

for $v - \delta_n < u < v$ and some $\delta_n > 0$. Hence

$$[fj + \chi_s]_u^v \geq f(v)[j]_u^v, \text{ and } [fj]_u^v + \chi_s(u)$$

is a major function in $[\alpha, v]$. Similarly

$$[fj]_u^v - \chi_s(u)$$

is a minor function in $[\alpha, v]$, and

$$[\chi_s]_u^v < 3\epsilon.$$

Thus $P(\alpha, v)$ exists. Similarly $P(v, \beta)$ exists, so that by (5, pp. 585-586), property I, $P(\alpha, \beta)$ exists, and v does not lie in T , contrary to hypothesis.

If j is of bounded variation in the closed interval J then $P(f, j; J)$ exists. Hence (45) is true. Further, if W contains an interval $[\xi, \eta]$ let J be a subinterval. If $P(f, j; J)$ exists for each bounded Baire function f , then by Theorem 1, and then Theorem 3(10), the set of points $\{d_n\}$ in J has the Denjoy property (see, e.g., (1), chap. III, p. 140). Hence it is scattered, and so is nowhere dense in J . It follows that W must be nowhere dense in J , as the points $\{u_n\}$ of Theorem 3 add nothing to W . This contradicts the fact that J is contained in W , so that $[\xi, \eta]$ is contained in T , and T contains the interior of W , proving (46).

To prove (47) we first take the closure J_n of the n th interval of the interior of Q , and construct a function j_n satisfying (37), (38) with $J = J_n$. Then we construct a function j_0 satisfying (37), (39), with the Q there replaced by the present Q less its interior. Finally we construct a function j_{-1} satisfying (37), (40), with the Q there replaced by the closure of $R - Q$. Then

$$\sum_{n=-1}^{\infty} j_n$$

satisfies the conditions of (47).

For (48), if T is empty then by Theorems 1 and 3(10), the set of points $\{d_n\}$ in $[a, b]$ has the Denjoy property, and so is scattered. But for the function satisfying (37), (39), the set of points $\{d_n\}$ in $[a, b]$ is also scattered, so that (48) follows.

Corollary 1 follows from (44), (45), and Corollary 2 from (47).

REFERENCES

1. A. Denjoy, *Leçons sur le calcul des coefficients d'une série trigonométrique* (Paris, 1941, 1949).
2. R. Henstock, *The efficiency of convergence factors for functions of a continuous real variable*, J. London Math. Soc., 30 (1955), 273-286.
3. J. E. Littlewood, *The Elements of the Theory of Real Functions* (Cambridge, 1926).
4. S. Saks, *Theory of the Integral* (Warsaw, 1937).
5. A. J. Ward, *The Perron-Stieltjes integral*, Math. Z., 41 (1936), 578-604.

Queen's University,
Belfast

A GENERALIZATION OF THE CAUCHY PRINCIPAL VALUE

CHARLES FOX

1. Introduction. If $a < u < b$ and $n > 0$ then

$$(1) \quad \int_a^b \frac{f(x) dx}{(x-u)^{n+1}}$$

is a so-called improper integral owing to the infinity in the integrand at $x = u$. When $n = 0$ we have associated with (1) the well-known Cauchy principal value, namely

$$(2) \quad \lim_{\epsilon \rightarrow 0} \left\{ \int_a^{u-\epsilon} \frac{f(x) dx}{(x-u)^{n+1}} + \int_{u+\epsilon}^b \frac{f(x) dx}{(x-u)^{n+1}} \right\}.$$

Hadamard (1, p. 117 et seq.) derives from an improper integral an expression which he calls its finite part and which, as he shows, possesses many important properties. For given $f(x)$ this finite part is obtained by constructing a function $g(x)$ so that the following limit exists (1, pp. 136 and 138):

$$(3) \quad \lim_{\epsilon \rightarrow 0} \int_a^b \frac{f(x) - g(x)}{(x-u)^m} dx.$$

Hadamard confines himself to the case when $m = n + \frac{1}{2}$ and n is a positive integer. In this paper we shall use Hadamard's idea to define a principal value for (1) in the case when n , in (1), is a positive integer greater than zero. When $n = 0$ the definition will reduce to (2).

The principal value so defined enables us to generalize several well-known theorems. We shall illustrate this generalization later by discussing the Hilbert transform (4, p. 120 (5.1.11)) and the Plemelj formulae (2, p. 42 (17.2)).

2. The principal value of (1). For the rest of this paper n will always denote a positive integer or zero; $f^i(x)$ will denote the i th derivative of $f(x)$ with respect to x and the principal value of (1) will be indicated by means of a prefix P before the integral sign.

For integration along the real axis with $a < u < b$ the principal value of (1) is defined as follows:

$$(4) \quad P \int_a^b \frac{f(x) dx}{(x-u)^{n+1}} = \lim_{\epsilon \rightarrow 0} \left\{ \int_a^{u-\epsilon} \frac{f(x) dx}{(x-u)^{n+1}} + \int_{u+\epsilon}^b \frac{f(x) dx}{(x-u)^{n+1}} - H_n(u, \epsilon) \right\}$$

where

Received March 14, 1956.

$$(5) \quad H_0(u, \epsilon) = 0, \quad n = 0,$$

$$(6) \quad H_n(u, \epsilon) = \sum_{i=0}^{n-1} \frac{f^{(i)}(u)}{i!} \left\{ \frac{1 - (-1)^{n-i}}{(n-i)\epsilon^{n-i}} \right\}, \quad n > 0.$$

When $n = 0$ this principal value evidently reduces to (2).

THEOREM 1. If (i) $f(x)$ possess derivatives up to order n in (a, b) and (ii) $f^n(x)$ satisfies a Lipschitz (or Hölder) condition, namely

$$(7) \quad |f^n(x_1) - f^n(x_2)| \leq A|x_1 - x_2|^\mu$$

whenever x_1 and x_2 both lie in (a, b) , A is a constant and $0 < \mu \leq 1$, then the limit on the right hand side of (4) exists.

Proof. Consider the expression E given by

$$(8) \quad E = \left\{ f(x) - f(u) - \frac{(x-u)}{1!} f'(u) - \dots - \frac{(x-u)^{n-1}}{(n-1)!} f^{n-1}(u) \right\} \frac{1}{(x-u)^{n+1}}.$$

From (i) and the mean value theorem, we see that

$$(9) \quad E = \frac{f^n(t)}{n! (x-u)},$$

where t lies between x and u . On writing

$$(10) \quad E = \left\{ \frac{f^n(t) - f^n(u)}{n! (x-u)} \right\} + \frac{f^n(u)}{n! (x-u)} = E_1 + E_2$$

we see from (ii) that, since t lies between x and u ,

$$(11) \quad |E_1| \leq A|x-u|^{\mu-1}, \quad -1 < \mu-1 \leq 0.$$

Consequently it follows from the usual theory of the Cauchy principal value (2, chap. 2) that

$$\lim_{\epsilon \rightarrow +0} \left\{ \int_a^{u-\epsilon} + \int_{u+\epsilon}^b \right\} E \, dx$$

exists.

Denote the part of (4) inside the brackets $\{\}$ by R . Then, apart from functions of u , we see that

$$(12) \quad R - \left\{ \int_a^{u-\epsilon} + \int_{u+\epsilon}^b \right\} E \, dx$$

contains only negative powers of $x-u$ in its integrand. On performing the integrations we find that (12) is independent of ϵ and it therefore follows that

$$\lim_{\epsilon \rightarrow +0} R$$

must also exist. This completes the proof of Theorem 1.

The definition (4) is easily extended to the case when the integration is taken along the arc of a plane curve. The variables are then all complex, a and b correspond to the end points of the arc of integration, x is any point on this arc and u is a fixed point on it. Draw a circle centre u and radius $\epsilon (>0)$ so small that the arc of integration is cut in two points only, $u - \epsilon_1$ between a and u and $u + \epsilon_2$ between u and b . The principal value is then obtained by making the following changes in the right hand side of (4): (i) in the first integral replace $u - \epsilon$ by $u - \epsilon_1$, (ii) in the second integral replace $u + \epsilon$ by $u + \epsilon_2$ and (iii) in $H_n(u, \epsilon)$ replace

$$(13) \quad \left\{ \frac{1 - (-1)^{n-i}}{(n-i)\epsilon^{n-i}} \right\} \text{ by } \frac{1}{(n-i)} \left\{ \frac{1}{\epsilon_1^{n-i}} - \frac{(-1)^{n-i}}{\epsilon_2^{n-i}} \right\}.$$

Theorem 1 holds for this definition also if the interval (a, b) is replaced by the arc of integration from a to b .

3. Some properties of the principal value (4).

THEOREM 2. *With the same conditions as in Theorem 1 and $0 < m < n$, $n \geq 1$, we have*

$$(14) \quad P \int_a^b \frac{f(x) dx}{(x-u)^{n+1}} = \sum_{i=0}^{n-1} \frac{(n-i-1)!}{n!} \left\{ \frac{f^i(a)}{(a-u)^{n-i}} - \frac{f^i(b)}{(b-u)^{n-i}} \right\} + \frac{(n-m)!}{n!} P \int_a^b \frac{f^m(x) dx}{(x-u)^{n-m+1}}.$$

Proof. Denote the part of (4) inside the brackets $\{\}$ by R . Denote the result of replacing $f(u)$ by $f^i(u)$ and n by $n-1$ in the right hand side of (6) by $K_{n-1}(u, \epsilon)$. On integrating the two integrals in R by parts we have

$$(15) \quad R = \frac{f(a)}{n(a-u)^n} - \frac{f(b)}{n(b-u)^n} - \frac{f(u-\epsilon)}{n(-\epsilon)^n} + \frac{f(u+\epsilon)}{n(\epsilon)^n} - H_n(u, \epsilon) + \frac{1}{n} K_{n-1}(u, \epsilon) + \frac{1}{n} \left\{ \int_a^{u-\epsilon} \frac{f^1(x) dx}{(x-u)^n} + \int_{u+\epsilon}^b \frac{f^1(x) dx}{(x-u)^n} - K_{n-1}(u, \epsilon) \right\}.$$

Denote the sum of the 3rd, 4th, 5th, and 6th terms on the right by S . From condition (i) we may expand $f(u-\epsilon)$ and $f(u+\epsilon)$, by the mean value theorem, in powers of ϵ as far as the n th derivative of $f(u)$. It is then easily found that all the coefficients of the various powers of ϵ vanish, so that S is equal to the remainder terms only. Consequently

$$(16) \quad S = \frac{-f''(t_1) + f''(t_2)}{n(n!)},$$

where t_1 lies between $u - \epsilon$ and u , while t_2 lies between u and $u + \epsilon$. From the Hölder condition it follows immediately that $S \rightarrow 0$ as $\epsilon \rightarrow +0$. On letting $\epsilon \rightarrow +0$ in (15) we see from (4) and the definition of $K_{n-1}(u, \epsilon)$ that

$$(17) \quad P \int_a^b \frac{f(x) dx}{(x-u)^{n+1}} = \frac{f(a)}{n(a-u)^n} - \frac{f(b)}{n(b-u)^n} + \frac{1}{n} P \int_a^b \frac{f^1(x) dx}{(x-u)^n}.$$

This is evidently (14) with $m = 1$. On applying (17) to the principal value on the right hand side of (17) we establish (14) for the case $m = 2$. By continuing this process m times we establish (14) for every integral value of $m \leq n$.

Theorem 2 is also true for complex integration:

THEOREM 2A. *If conditions (i) and (ii) of Theorem 1 hold with $a = -\infty$ and $b = +\infty$, (iii) for large $|x|$, $f^m(x) = O(x^{n-m-p})$ ($p > 0$, $m \leq n$), where O is the Landau order symbol and (iv) $f^i(x)/x^{n-i} \rightarrow 0$ ($i = 0, 1, \dots, m-1$) when $x \rightarrow \infty$ or $x \rightarrow -\infty$, then*

$$(18) \quad P \int_{-\infty}^{\infty} \frac{f(x) dx}{(x-u)^{n+1}} = \frac{(n-m)!}{n!} P \int_{-\infty}^{\infty} \frac{f^m(x) dx}{(x-u)^{n-m+1}} \quad (m \leq n).$$

Proof. From (iii) the integrals in (14) converge when $a = -\infty$ and $b = +\infty$ and so, from (i) and (ii), (14) is true with $-\infty$ and ∞ as the limits of integration. From (iv) all the terms in the summation sign of (14) vanish for such infinite limits and so (14) reduces to (18).

THEOREM 2B. *If $f(z)$ is one valued and analytic in a domain which includes the simple closed Jordan curve C and its interior then*

$$(19) \quad P \int_C \frac{f(z) dz}{(z-u)^{n+1}} = \frac{(n-m)!}{n!} P \int_C \frac{f^m(z) dz}{(z-u)^{n-m+1}} \quad (m \leq n),$$

where the integrals are taken once round C and u is a fixed point on C .

Proof. Since $f(z)$ is analytic it possesses derivatives of all orders, each derivative satisfying a Lipschitz condition. Hence (14), with integration along the contour C , is true when $f(z)$ is analytic. Since C is closed the end points coincide, i.e. $b = a$. Hence since $f(z)$ and its derivatives are one valued it follows that the terms in the summation sign of (14) cancel in pairs, leaving us with (19).

4. An extension of the Hilbert transform.

THEOREM 3. *If the conditions of Theorem 2A hold, and (v) for large $|x|$ $f^p(x) = O(x^{-1-p})$ for $p > 0$, and*

$$(20) \quad g(u) = \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{f(x) dx}{(x-u)^{n+1}},$$

where u is real, then

$$(20.1) \quad g(x) \in L^2(-\infty, \infty):$$

$$(21) \quad f^n(u) = -\frac{n!}{\pi} P \int_{-\infty}^{\infty} \frac{g(x) dx}{(x-u)^n},$$

and

$$(22) \quad \int_{-\infty}^{\infty} \{g(x)\}^2 dx = \frac{1}{(n!)^2} \int_{-\infty}^{\infty} \{f^n(x)\}^2 dx.$$

Proof. Since Theorem 2A holds we may use (18). On using (18) with $m = n$ it follows that (20) can be written in the form

$$(23) \quad g(u) = \frac{1}{\pi(n!)} P \int_{-\infty}^{\infty} \frac{f^n(x) dx}{(x-u)^n}.$$

On replacing x by $u-t$ in the range $-\infty < x < u-\epsilon$ and x by $u+t$ in the range $u+\epsilon < x < \infty$, (23) becomes

$$(24) \quad g(u) = \frac{1}{\pi(n!)} \lim_{\epsilon \rightarrow +0} \int_{\epsilon}^{\infty} \frac{f^n(u+t) - f^n(u-t)}{t} dt.$$

From (v) $f^n(x) \in L^2(-\infty, \infty)$ and so we may apply Hilbert's transform theorem (4, Theorem 91, p. 122) to (24). The truth of (20.1), (21) and (22), including the existence of the principal value on the right of (21), then follows immediately.

Evidently we may look upon (20) as an integral equation with $g(x)$ as a known and $f(x)$ as an unknown function. The solution is then given by (21). Theorem 3 can also be established under a different set of conditions if we use M. Riesz's version of the Hilbert Transform (4, p. 132).

5. An extension of the Plemelj formulae. Let A denote an arc in the complex z plane generated by points $z = x + iy$ where $x = x(t)$ and $y = y(t)$ are continuous single valued functions of the real variable t . We shall assume that there is a unique tangent at each point of the arc and we shall denote the end points by E_1 and E_2 . On describing A from E_1 to E_2 we can divide the neighbourhood of each point u on A (other than E_1 and E_2) into two areas, a left hand area and a right hand area with a small segment of A , containing u , as a boundary between the two areas. Certain functions $h(v)$ possess the following property. Let u be a point on the arc A , other than one of the end points, then as $v \rightarrow u$ the function $h(v)$ tends to a unique limit provided that the path to u lies *entirely* either in the left hand area or in the right hand area. If this is the case then the limit from the left hand area will be denoted by $h^+(u)$ and that from the right hand area by $h^-(u)$. We exclude paths which approach u ultimately along the tangent to A .

When A becomes a closed contour C then C is described so that its exterior, containing the point at infinity, is the right hand area.

Let $F(v)$ be defined by the equation

$$(25) \quad F(v) = \frac{1}{2\pi i} \int_A \frac{f(z) dz}{(z-v)^{n+1}},$$

where the integration is taken along the arc A .

THEOREM 4. If (i) for all points of A , except possibly the end points, $f(z)$ possess derivatives up to order n , (ii) $f(z)$ satisfies a Hölder condition on A

(see inequality (7)) and u is a point on A other than one of the end points, then $F^+(u)$ and $F^-(u)$ both exist and

$$(26) \quad F^+(u) = \frac{1}{2(n!)} f^n(u) + \frac{1}{2\pi i} P \int_A \frac{f(z) dz}{(z-u)^{n+1}},$$

$$(27) \quad F^-(u) = -\frac{1}{2(n!)} f^n(u) + \frac{1}{2\pi i} P \int_A \frac{f(z) dz}{(z-u)^{n+1}}.$$

Proof. With v not a point on A integrate (25) n times by parts. We obtain

$$(28) \quad F(v) = G(v) + \frac{1}{2\pi i(n!)} \int_A \frac{f^n(z) dz}{(z-v)},$$

where

$$(29) \quad G(v) = \frac{1}{2\pi i} \sum_{i=0}^{n-1} \frac{(n-i-1)!}{n!} \left\{ \frac{f^i(a)}{(a-v)^{n-i}} - \frac{f^i(b)}{(b-v)^{n-i}} \right\}.$$

The conditions assumed above ensure the truth of Theorems 1 and 2 for the case of integration along the arc A . Hence when $v = u$, where u is a point on A , the integral on the right of (25) has a principal value. Again on taking the case when $m = n$ in (14), multiplying by $1/(2\pi i)$ and subtracting from (28) we obtain

$$(30) \quad F(v) - \frac{1}{2\pi i} P \int_A \frac{f(z) dz}{(z-u)^{n+1}} \\ = G(v) - G(u) + \frac{1}{2\pi i(n!)} \left\{ \int_A \frac{f^n(z) dz}{(z-v)} - P \int_A \frac{f^n(z) dz}{(z-u)} \right\}.$$

Now let

$$(31) \quad h(v) = \frac{1}{2\pi i} \int_A \frac{f^n(z) dz}{(z-v)}.$$

Then if u is a point on A , since $f^n(z)$ satisfies a Hölder condition, it is known that $h^+(u)$ and $h^-(u)$ exist (2, §16, p. 37). Again as $v \rightarrow u$ we have $\{G(v) - G(u)\} \rightarrow 0$. Hence on making $v \rightarrow u$ through the left hand area it follows from (30) that $F^+(u)$ exists and that

$$(32) \quad F^+(u) - \frac{1}{2\pi i} P \int_A \frac{f(z) dz}{(z-u)^{n+1}} = \frac{1}{n!} \{h^+(u) - \frac{1}{2\pi i} P \int_A \frac{f^n(z) dz}{(z-u)}\}.$$

From the first of the Plemelj formulae given by Muskhelishvili (2, p. 42 (17.2)) we see that the right hand side of (32) is equal to $f^n(u)/2(n!)$ which establishes the truth of (26).

Similarly, by using the second of the Plemelj formulae just cited we can establish the fact that $F^-(u)$ exists and also the truth of (27).

When we place $n = 0$ in (26) and (27) they reduce to the Plemelj formulae.

The theorem still holds if A is a simple closed contour. It can be extended to the case when the path of integration is the real axis, from $-\infty$ to ∞ , if suitable conditions are imposed upon the derivatives of $f(x)$ in order to make the integrals converge, for example conditions (iii) and (iv) of Theorem 2A.

We now see that $F(v)$, as defined in (25) with v complex, possesses the following properties: (i) if v is not on A it is an analytic function of v , (ii) for large v it is $O(v^{-n-1})$ and (iii) the arc A is a line of discontinuity. In fact from (26) and (27), when u is on A we have

$$(33) \quad F^+(u) - F^-(u) = \frac{1}{(n!)} f^n(u).$$

Again, with u on A , $F(u)$ is undefined, but if the conditions of Theorem 4 are satisfied we may define $F(u)$ to be equal to the principal value of the integral on the right of (25).

If a is one of the end points of A and $f(z)$ has a zero of order r at $z = a$ then it is not difficult to see that for $r > n$ $F(a)$ exists and that $F^+(a) = F^-(a) = F(a)$. If $r \leq n$ then in general $F(v)$ has a singularity at $v = a$ which is the sum of a logarithmic singularity and a pole.

6. Two applications of (33). When $n = 0$, (33) reduces to a result which can be derived from the Cauchy principal value, a result which can be used to solve many important boundary problems in various branches of mathematical physics (2, chaps. 12 and 13; 3 pt. V). We now discuss briefly two such problems where we can use (33) in the more general case when $n > 0$ (n an integer).

Problem 1. Find a function $F(v)$ which (i) is analytic at all points v except for points on the arc A , (ii) for large v is $O(v^{-n-1})$ and (iii) for given $g(u)$, where u is on the arc A but is not one of its end points, we have

$$(34) \quad F^+(u) - F^-(u) = g(u).$$

To obtain a formal solution we first solve the differential equation

$$(35) \quad f^n(u) = (n!) g(u)$$

for $f(u)$ and then by an obvious substitution we can express (34) in the form (33). We then obtain as a formal solution of our problem

$$(36) \quad F(v) = \frac{1}{2\pi i} \int_A \frac{f(z) dz}{(z-v)^{n+1}}.$$

If $g(z)$ is an analytic function of z in a domain D which includes the arc A then there exists a solution of (35), $f(z)$ say, which is also analytic in D . Since $f(z)$ then satisfies the conditions of Theorem 4 it will follow that $F(v)$ is a solution of the problem. With a more prolonged discussion it is possible to show that a solution exists if $g(z)$ satisfies a Hölder condition along the arc A .

Problem 2. This is connected with the singular integral equation. For the case $n = 0$ the singular integral equation below has been studied in great detail by Muskhelishvili (2, chap. 6). For the general case $n > 1$ many new

difficulties occur but in one case a solution can be obtained by means of a reduction to Problem 1. The equation in question is

$$(37) \quad g(u) f^n(u) + \frac{h(u)}{\pi i} P \int_A \frac{f(z) dz}{(z-u)^{n+1}} = k(u),$$

where $g(u)$, $h(u)$ and $k(u)$ are given along the arc A and $f(z)$ is to be determined.

A formal solution is obtained by assuming that a function $F(v)$ exists which is related to $f(z)$ as in (25) and for which (26) and (27) both hold. On adding and subtracting (26) and (27) we obtain both $f^n(u)$ and the integral in (37) in terms of $F^+(u)$ and $F^-(u)$. After an obvious division (37) is then transformed to

$$(38) \quad F^+(u) = \left\{ \frac{g(u) (n!) - h(u)}{g(u) (n!) + h(u)} \right\} F^-(u) + \left\{ \frac{k(u)}{g(u) (n!) + h(u)} \right\}.$$

This equation can be reduced to the same type of equation as is solved in Problem 1, namely (34) with functions transformed from $F^+(u)$, $F^-(u)$, $g(u)$, $h(u)$ and $k(u)$ of (38) by means of known operations. $F(v)$ can then be found and then, by using (33), $f(z)$ can be found by integration.

The most important part of this solution is the reduction of (38) to an equation of type (34), an equation which is solvable by means of the methods of Problem 1. This reduction does not depend upon the value of n and is therefore the same for the general value of n as for the case when $n = 0$. The details and the ingenious methods used by Muskhelishvili to effect this reduction when $n = 0$ can be found in (2, §47, p. 123). If the coefficient of $F^-(u)$ in (38) and the second term on the right hand side of (38) both satisfy Hölder conditions then the solution obtained by this method is valid.

REFERENCES

1. J. Hadamard, *Lectures on Cauchy's Problem in Linear Partial Differential Equations* (New York, 1952).
2. N. I. Muskhelishvili, *Singular Integral Equations*, Translated from Russian into English by J. R. M. Radok (Groningen, 1953).
3. ———, *Some Basic Problems of the Mathematical Theory of Elasticity*, Translated from Russian into English by J. R. M. Radok (Groningen, 1953).
4. E. C. Titchmarsh, *An Introduction to the Theory of Fourier Integrals* (Oxford, 1937).

McGill University

ON THE ZEROS OF THE FRESNEL INTEGRALS

ERWIN KREYSZIG

1. Introduction. This paper is concerned with the Fresnel integrals

$$(1.1) \quad C(u) = \int_0^u \cos(p^2) dp, \quad S(u) = \int_0^u \sin(p^2) dp$$

in the complex domain.

Recent research work in different fields of physical and technical applications of mathematics shows that an increasing number of problems require a detailed knowledge of elementary and higher functions for complex values of the argument. The Fresnel integrals, introduced by A. J. Fresnel (1788-1827) in connection with diffraction problems, are among these functions; a small collection of papers of the above-mentioned kind is included in the bibliography at the end of this paper (3; 5; 7; 12; 13; 17; 19; 20; 22). Moreover, the Fresnel integrals are important since various types of more complicated integrals can be reduced (6) to analytic expressions involving $C(u)$ and $S(u)$.

At the present time the detailed investigation of special functions for complex argument is still in its infancy. It has been limited, until now, to certain classes of functions, especially to those which have the advantage of possessing simple functional relations or of satisfying an ordinary differential equation; in the latter case the theory of differential equations can be used when considering these functions. The Gamma function and the Bessel functions, respectively, are of this kind. The Fresnel integrals do not possess these advantageous properties and must therefore be treated by other methods.

The Fresnel integrals have been considered from different points of view (1; 2; 4; 8; 9; 14; 16; 18), but, until a short time ago, for a real argument only. The first two investigations (10; 11) of these functions for complex values of the argument include some initial results about the zeros and also two small tables of function values.

In this paper we shall prove some lemmas which yield a much more refined knowledge of the two integrals under consideration. Furthermore, we shall indicate relations to other known functions and develop new methods for investigating and computing the zeros of these integrals. We shall find large domains of the complex plane which cannot contain a zero of the Fresnel integrals. When determining the position of the zeros of a function it is always important to find (more or less accurate) approximate values for those zeros; then the computation up to the desired degree of accuracy can be done

Received May 12, 1956.

This work was supported by the National Research Council of Canada.

schematically by means of the usual iterative methods. We shall see that in the case of the Fresnel integrals such approximate values (which are even of great accuracy) can be obtained in a simple manner. Also the more exact determination of the zeros will turn out to be relatively easy if appropriate representations of these functions are used. A table of the values of some zeros of the Fresnel integrals can be found at the end of the paper.

2. Fundamental relations, asymptotic behaviour. It is advantageous to transform the integrals (1.1) by means of the substitution $p^2 = t$. In this manner we obtain

$$C(u) = \frac{1}{2} \int_0^{u^2} t^{-1/2} \cos t \, dt, \quad S(u) = \frac{1}{2} \int_0^{u^2} t^{-1/2} \sin t \, dt.$$

Since we will primarily investigate the zeros of those functions the factor $\frac{1}{2}$ becomes unessential, and we will therefore omit it. We write

$$(2.1) \quad C(z) = \int_0^z t^{-1/2} \cos t \, dt, \quad S(z) = \int_0^z t^{-1/2} \sin t \, dt$$

where $z = x + iy$ denotes a complex variable. The representations (2.1) will be used in what follows.

For finite values of $|z|$ the Taylor series development of the integrands of (2.1) at $z = 0$ may be integrated term by term. We find

$$(2.2) \quad C(z) = z^{1/2} \sum_{m=0}^{\infty} \frac{(-1)^m}{(2m)! (2m + \frac{1}{2})} z^{2m},$$

$$S(z) = z^{1/2} \sum_{m=0}^{\infty} \frac{(-1)^m}{(2m+1)! (2m + \frac{3}{2})} z^{2m+1}.$$

The functions $z^{-1/2}C(z)$ and $z^{-1/2}S(z)$ are entire transcendental functions. $C(z)$ and $S(z)$ have a branch point at $z = 0$; from (2.2) we obtain the relations

$$(2.3) \quad C(z e^{ik\pi}) = e^{ik\pi/2} C(z), \quad S(z e^{ik\pi}) = e^{-ik\pi/2} S(z).$$

Furthermore,

$$(2.4) \quad C(\bar{z}) = \overline{C(z)}, \quad S(\bar{z}) = \overline{S(z)}.$$

Hence we may limit our considerations to the first quadrant ($x > 0$, $y > 0$) of the z -plane.

We now consider the asymptotic behaviour of the Fresnel integrals. In order that the limits

$$(2.5) \quad C = \lim_{z \rightarrow \infty} \int_0^z t^{-1/2} \cos t \, dt, \quad S = \lim_{z \rightarrow \infty} \int_0^z t^{-1/2} \sin t \, dt$$

exist, we must choose a path of integration which goes asymptotically parallel to the real axis ($y = 0$) to infinity. Then C and S have a uniquely determined finite value; transforming Euler's integral representation of the Gamma function in a suitable manner we find

$$(2.6) \quad C = S = 2^{-1} \Gamma(\frac{1}{2}) = \sqrt{(\pi/2)} = 1.253 \, 314 \, 1 \dots$$

Using such a path of integration, we have

$$(2.7) \quad C(z) + c(z) = C, \quad S(z) + s(z) = S,$$

where

$$(2.8) \quad c(z) = \int_z^\infty t^{-1} \cos t \, dt, \quad s(z) = \int_z^\infty t^{-1} \sin t \, dt.$$

Integrating (2.8) by parts and using (2.7), we obtain the following series representation of the Fresnel integrals:

$$(2.9) \quad \begin{aligned} (a) \quad C(z) &\sim C + z^{-1/2}(a(z) \cos z + b(z) \sin z), \\ (b) \quad S(z) &\sim S + z^{-1/2}(-b(z) \cos z + a(z) \sin z) \end{aligned}$$

where

$$a(z) = \sum_{m=1}^{\infty} (-1)^m \frac{1.3 \dots (4m-3)}{(2z)^{2m-1}}, \quad b(z) = 1 + \sum_{m=1}^{\infty} (-1)^m \frac{1.3 \dots (4m-1)}{(2z)^{2m}}.$$

LEMMA 1. The series (2.9) are asymptotic expansions of $C(z)$ and $S(z)$, respectively, for all complex values of z with the exception of pure imaginary ones.

Proof. Let us assume that we have integrated (2.7) by parts a number of times so that the integrated term finally obtained involves the power $z^{-m+1/2}$. Then, when neglecting some numerical factor which does not interest us, the remaining integral is of the form

$$\int_z^\infty t^{-(m+1/2)} \cos t \, dt \quad \text{or} \quad \int_z^\infty t^{-(m+1/2)} \sin t \, dt.$$

We thus have to estimate the integrals

$$K_1 = \int_z^\infty t^{-(m+1/2)} e^{it} \, dt, \quad K_2 = \int_z^\infty t^{-(m+1/2)} e^{-it} \, dt.$$

Setting $t = z + iw$ and $t = z - iw$, respectively, we obtain

$$\begin{aligned} K_1 &= ie^{iz} z^{-(m+1/2)} \int_0^\infty (1 + iw/z)^{-(m+1/2)} e^{-w} \, dw, \\ K_2 &= -ie^{-iz} z^{-(m+1/2)} \int_0^\infty (1 - iw/z)^{-(m+1/2)} e^{-w} \, dw. \end{aligned}$$

If z is not a pure imaginary number, i.e. $|\arg z| < \frac{1}{2}\pi - \gamma$ or $|\arg z| > \frac{1}{2}\pi + \gamma$, $\gamma > 0$, the inequality

$$|1 \pm iw/z| \geq \sin \gamma$$

holds, and therefore

$$\left| \int_0^\infty (1 \pm iw/z)^{-(m+1/2)} e^{-w} \, dw \right| < (\operatorname{cosec} \gamma)^{m+1/2}.$$

Hence

$$e^{-iz} K_1 = O(z^{-(m+1/2)}), \quad e^{iz} K_2 = O(z^{-(m+1/2)}).$$

The constant involved in the Landau symbol does not depend on $\arg z$ but depends on γ and tends to infinity if γ tends to zero.

A value of z being given, the greatest possible accuracy is obtained if the number of terms of (2.9) is chosen so that the last term corresponds to the highest value of m for which

$$(2.10) \quad m \leq \frac{1}{2} \{ (|z|^2 + \frac{1}{4})^{\frac{1}{2}} + 1 \}$$

still holds, as can easily be seen.

3. Relations to other known functions. The relations of the Fresnel integrals to the incomplete gamma functions

$$(3.1) \quad (a) P(\phi, z) = \int_0^z e^{-t} t^{\phi-1} dt, \quad (b) Q(\phi, z) = \int_z^\infty e^{-t} t^{\phi-1} dt = \Gamma(\phi) - P(\phi, z)$$

are of basic importance. Setting $\phi = \frac{1}{2}$ and substituting $t = iw$ and $t = -iw$, respectively, we obtain from (3.1a)

$$(3.2) \quad \begin{aligned} C(z) &= \frac{1}{2} [i^{-\frac{1}{2}} P(\frac{1}{2}, iz) + i^{\frac{1}{2}} P(\frac{1}{2}, -iz)], \\ S(z) &= \frac{1}{2} [i^{\frac{1}{2}} P(\frac{1}{2}, iz) + i^{-\frac{1}{2}} P(\frac{1}{2}, -iz)]. \end{aligned}$$

When computing a table of a special function the situation is very often as follows: For small arguments the Taylor series development at $z = 0$ can be used and for large values of $|z|$ the asymptotic expansion permits a simple calculation. The remaining difficulty consists in determining function values for arguments which are not very close to $z = 0$ but are too small to be calculated exactly enough by means of the asymptotic expansion. With respect to the Fresnel integrals we are just in such a situation, but we can overcome the difficulty by using the Nielsen representation (16, p. 84):

$$(3.3) \quad Q(\phi, z + h) = Q(\phi, z) - e^{-z} \sum_{m=0}^{\infty} (-1)^m \binom{m-\phi}{\phi} \frac{P(m+1, h)}{z^{m+1-\phi}}.$$

Setting $\phi = \frac{1}{2}$, $t = iw$, and $t = -iw$, respectively we obtain from (3.1b),

$$(3.4) \quad \begin{aligned} c(z) &= \frac{1}{2} [i^{-\frac{1}{2}} Q(\frac{1}{2}, iz) + i^{\frac{1}{2}} Q(\frac{1}{2}, -iz)], \\ s(z) &= \frac{1}{2} [i^{\frac{1}{2}} Q(\frac{1}{2}, iz) + i^{-\frac{1}{2}} Q(\frac{1}{2}, -iz)], \end{aligned}$$

and from this and (3.3),

$$(3.5) \quad \begin{aligned} c(z+h) &= c(z) + i2^{-1} z^{-\frac{1}{2}} (P_1(z) - P_2(z)), \\ s(z+h) &= s(z) - 2^{-1} z^{-\frac{1}{2}} (P_1(z) + P_2(z)) \end{aligned}$$

where

$$(3.6) \quad \begin{aligned} P_1(z) &= e^{-iz} \sum_{m=0}^{\infty} i^m \binom{m-\frac{1}{2}}{\frac{1}{2}} \frac{P(m+1, ih)}{z^m}, \\ P_2(z) &= e^{iz} \sum_{m=0}^{\infty} i^{-m} \binom{m-\frac{1}{2}}{\frac{1}{2}} \frac{P(m+1, -ih)}{z^m}. \end{aligned}$$

By means of (2.8) and (3.5) we are able to calculate the first differences of the function values of $C(z)$ and $S(z)$. Starting then from function values which can be simply obtained by using (2.2) or (2.9) we can immediately compute the desired function values of the Fresnel integrals. For the above-mentioned "medium" values of the argument this procedure is much better than a direct calculation by means of (2.2). From (3.1) we have

$$(3.7) \quad (a) P(1, h) = 1 - e^{-h}, \quad (b) P(m+1, h) = \int_0^h e^{-t} t^m dt.$$

Starting from (3.7a) and using the recurrence relation

$$P(m+1, h) = mP(m, h) - e^{-h} h^m,$$

the functions $P(m+1, ih)$ and $P(m+1, -ih)$ occurring in (3.6) can be easily calculated. It is advantageous, of course, to choose a fixed value of h for a certain computation.

For the sake of completeness we mention also the following relations: using the integral representation (15, p. 87)

$${}_1F_1(a, c; z) = \frac{\Gamma(c)}{\Gamma(a) \Gamma(c-a)} z^{1-c} \int_0^z e^{t^{a-1}} (z-t)^{c-a-1} dt$$

of Kummer's function ${}_1F_1(a, c; z)$, setting $a = \frac{1}{2}$, $c = \frac{3}{2}$ and $t = \pm iw$, respectively, we obtain

$$(3.8) \quad \begin{aligned} C(z) &= \sqrt{z} [{}_1F_1(\tfrac{1}{2}, \tfrac{3}{2}; -iz) + {}_1F_1(\tfrac{1}{2}, \tfrac{3}{2}; iz)] \\ S(z) &= i\sqrt{z} [{}_1F_1(\tfrac{1}{2}, \tfrac{3}{2}; -iz) - {}_1F_1(\tfrac{1}{2}, \tfrac{3}{2}; iz)]. \end{aligned}$$

It is known that the Fresnel integrals are also related to the error function

$$\phi(z) = \frac{2}{\pi} \int_0^z e^{-t^2} dt;$$

substituting $t = \sqrt{(\pm iw)}$ we have

$$(3.9) \quad \begin{aligned} C(z) &= \tfrac{1}{2}\pi [i^{\frac{1}{2}}\phi(\sqrt{-iz}) + i^{-\frac{1}{2}}\phi(\sqrt{iz})] \\ S(z) &= \tfrac{1}{2}\pi [i^{-\frac{1}{2}}\phi(\sqrt{-iz}) + i^{\frac{1}{2}}\phi(\sqrt{iz})]. \end{aligned}$$

4. Domains which cannot contain zeros. Let us first give a simple proof of the fact that all zeros $z (\neq 0)$ of the Fresnel integrals must be complex.

LEMMA 2. *The Fresnel integrals do not vanish for any real or purely imaginary argument different from zero. All zeros of these functions are simple and conjugate complex to each other in pairs.*

Proof. In consequence of (2.3) we may consider positive values of x only. From the form of the integrand of (2.1) it follows that

$$(4.1) \quad \left. \begin{aligned} (a) \quad & C((4n+1)\pi/2) - C((4n-1)\pi/2) > 0, \\ & C((4n+3)\pi/2) - C((4n+1)\pi/2) < 0, \\ (b) \quad & S((2n+1)\pi) - S(2n\pi) > 0, \\ & S(2n\pi) - S((2n-1)\pi) < 0, \end{aligned} \right\} \begin{aligned} n &= 1, 2, \dots, \\ n &= 0, 1, \dots, \\ n &= 1, 2, \dots, \end{aligned}$$

Since $x^{-1/2}$ is monotone,

$$(a) \quad |C((2n+3)\pi/2) - C((2n+1)\pi/2)| < |C((2n+1)\pi/2) - C((2n-1)\pi/2)|$$

$$(b) \quad |S((n+1)\pi) - S(n\pi)| < |S(n\pi) - S((n-1)\pi)|, \quad n = 1, 2, \dots$$

From (4.1b), (4.2b), and $S(0) = 0$ it follows that $S(x) \neq 0$ for any real value of $x \neq 0$. In order to draw the same conclusion with respect to $C(x)$ from (4.1a) and (4.2a) we have to prove that $C(3\pi/2) > 0$. Using (2.6), (2.8) and integrating by parts, we find

$$C(3\pi/2) = \sqrt{(\pi/2)} - \sqrt{(2/3\pi)} + \frac{3}{4} \int_{3\pi/2}^{\infty} t^{-5/2} \cos t \, dt.$$

Since $t^{-5/2}$ is monotone,

$$\left| \int_{(2n-1)\pi/2}^{(2n+1)\pi/2} t^{-5/2} \cos t \, dt \right| > \left| \int_{(2n+1)\pi/2}^{(2n+3)\pi/2} t^{-5/2} \cos t \, dt \right|, \quad n = 1, 2, \dots$$

Consequently

$$\int_{3\pi/2}^{\infty} t^{-5/2} \cos t \, dt > 0.$$

Hence $C(3\pi/2) > 0$. This completes the proof that $C(z)$ cannot vanish for real values of $x (\neq 0)$. If z is a pure imaginary number all terms in (2.2) have the same sign; hence there cannot exist a pure imaginary zero of the Fresnel integrals. The existence of zeros follows from the fact that $z^{1/2}C(z)$ and $z^{1/2}S(z)$ are entire functions which are not of the kind of an exponential function. Since the integrands of the Fresnel integral have real zeros only the zeros of the integrals are simple. Since (2.2) has real coefficients the zeros of the Fresnel integrals are conjugate complex in pairs. This completes the proof of Lemma 2.

We now consider the possibility of limiting the zeros of $C(z)$ and $S(z)$ to certain domains of the complex plane.

THEOREM 3. *The Fresnel integral $C(z)$ cannot possess zeros in any of the strips which are parallel to the y -axis and correspond to the values*

$$0 < x \leq \pi, \quad (4n-1)\pi/2 \leq x \leq (2n+1)\pi, \quad n = 1, 2, \dots$$

The same is true for $S(z)$ with respect to the strips parallel to the y -axis and corresponding to the values

$$0 < x \leq 3\pi/2, \quad 2n\pi \leq x \leq (4n+3)\pi/2, \quad n = 1, 2, \dots$$

Proof. Because of (2.3) and (2.4) we may consider the first quadrant ($x < 0$, $y < 0$) of the z -plane only. In order to prove the first of the two statements we start from the integral

$$J = C(x+iy) - C(x) = \int_x^{x+iy} t^{-1/2} \cos t \, dt.$$

Setting $t = x + iw$ and $(x + iw)^{-1} = a + ib$ we obtain

$$\Re J = \sin x \int_0^y a \sinh w \, dw - \cos x \int_0^y b \cosh w \, dw,$$

$$\Im J = \sin x \int_0^y b \sinh w \, dw + \cos x \int_0^y a \cosh w \, dw.$$

Since $x > 0$ and $y > 0$ we have $a > 0$ and $b < 0$ for all values of z under consideration. $C(x)$ is real and not negative, cf. Lemma 2. We thus obtain

$$\Re J > 0, \Re C(z) > 0 \quad (2n\pi < x < (4n+1)\pi/2),$$

and

$$\Im J < 0 \quad ((4n+1)\pi/2 < x < (2n+1)\pi),$$

$$\Im J > 0 \quad ((4n+3)\pi/2 < x < (2n+2)\pi).$$

From this the statement on $C(z)$ follows. The second part of Theorem 3 can be proved in a similar manner.

It should be noticed that the idea of the proof of Theorem 3 can be applied to more general integrals of the type

$$(4.3) \quad C(z, \alpha) = \int_0^z t^{-\alpha} \cos t \, dt, \quad 0 < \alpha < 1$$

in order to obtain the same result on the zeros. Also the integrals

$$(4.4) \quad S(z, \alpha) = \int_0^z t^{-\alpha} \sin t \, dt, \quad 0 < \alpha < 1$$

may be considered in this manner, but the method is not applicable to integrals (4.4) having values of α between 1 and 2 (exclusively), since in this case $\Re t^{-\alpha} > 0$ and $\Im t^{-\alpha} < 0$ may not hold. Indeed, for sufficiently large values of α (< 2), $S(z, \alpha)$ has zeros also outside of the strips defined by

$$(4n-1)\pi/2 < x < 2n\pi.$$

5. Formulas of approximation for the zeros. From Theorem 3 we can draw the important conclusion that all zeros of $C(z)$ and $S(z)$ are at a sufficiently large distance from the origin $z = 0$. This fact enables us to use the asymptotic expansion (2.9) for a more detailed investigation of those zeros.

As was pointed out in the introduction, it is always important to have approximation formulas for the position of the zeros of a function, since approximate values can yield the starting point for applying one of the usual iterative methods for a more accurate determination of those zeros. We will now derive simple approximation formulas for the zeros of $C(z)$ and $S(z)$.

In consequence of (2.9) the zeros of the equation

$$(5.1) \quad \sin z = -C\sqrt{z}$$

are first approximations of the zeros of $C(z)$. Setting $\sqrt{z} = p + iq$ and using (2.6), we obtain from (5.1)

$$(5.2) \quad p = -\sqrt{(2/\pi)} \sin x \cosh y, \quad q = -\sqrt{(2/\pi)} \cos x \sinh y.$$

We consider the strip S_n : $(2n-2)\pi < x < 2n\pi$, $y > 0$, which is parallel to the y -axis. In consequence of Theorem 3, only the part $S_n' \subset S_n$, defined by $(2n-1)\pi < x < (4n-1)\pi/2$, $y > 0$ can contain a complex zero of $C(z)$. Setting

$$H(z, \alpha) = \int_0^z t^{-\alpha} \cos t \, dt, \quad \frac{1}{2} > \alpha > 0$$

we have $H(z, 0) = \sin z$ and $H(z, \frac{1}{2}) = C(z)$. $H(z, 0)$ has real zeros and $H(z, \frac{1}{2})$ has complex zeros only. Hence, if α decreases monotonely from $\frac{1}{2}$ to 0 then, for a certain value $\alpha_0 = \alpha_0(n)$, $\frac{1}{2} > \alpha_0 > 0$ we must have a real zero z_{n_0} of $H(z, \alpha)$ in S_n a first time. Since, for all values of α , $H(z, \alpha)$ has a minimum at $x = (4n-1)\pi/2$ the zero z_{n_0} must coincide with that point. Hence, when denoting by x_n the real part of the zero of $C(z)$ in S_n' and setting

$$(5.3) \quad x_n = (4n-1)\pi/2 - \gamma_n,$$

because of continuity γ_n must be a small (positive) quantity. We have

$$\cos x_n = -\gamma_n + O(\gamma_n^3), \quad \sin x_n = -1 + O(\gamma_n^2),$$

where the functions indicated by the Landau symbol are small of higher order. The absolute values of the zeros of $C(z)$ —even that of the smallest one—are relatively large, cf. Theorem 3. Hence the same is true for the corresponding quantities $|p|$. Setting

$$\cosh y = \frac{1}{2}e^y + O(e^{-y}), \quad \sinh y = \frac{1}{2}e^y + O(e^{-y}),$$

the second term is thus small in comparison with the first one. Omitting the functions indicated by the Landau symbols and using

$$(5.4) \quad x = p^2 - q^2 = \frac{2}{\pi}(\sin^2 x \cosh^2 y - \cos^2 x \sinh^2 y),$$

cf. (5.2), we obtain the following approximate expression $y_{0,n}$ for the imaginary part y_n of the zero of $C(z)$ which is located in S_n :

$$(5.5) \quad y_{0,n} = \log(\pi\sqrt{(4n-1)}), \quad n = 1, 2, \dots$$

From (5.3)–(5.5) and

$$y = 2pq = \frac{1}{\pi} \sin 2x \sinh 2y,$$

cf. (5.2), we obtain a similar approximation $x_{0,n}$ for the real part x_n of the zero of $C(z)$ which is located in S_n :

$$(5.6) \quad x_{0,n} = \frac{(4n-1)\pi}{2} - \frac{\log(\pi\sqrt{(4n-1)})}{(4n-1)\pi}, \quad n = 1, 2, \dots$$

The degree of accuracy of these simple formulas (5.5) and (5.6) is relatively high. For the real and imaginary part of the smallest zero of $C(z)$ the error amounts to 2 per cent and 1 per cent, respectively. The accuracy increases rapidly with increasing values of $|z|$, as can easily be proved.

Comparing (5.5) and (5.6) we find

LEMMA 4. *The imaginary part y_n of the n th zero of $C(z)$ increases monotonely with n . We have*

$$(5.7) \quad \lim_{n \rightarrow \infty} y_n/x_n = 0.$$

The difference γ_n between x_n and $(4n-1)\pi/2$ decreases monotonely with increasing n and tends to zero if n tends to infinity.

Using (5.1), the third term of (2.9) takes the form

$$\frac{1}{2} z^{-3/2} \cos z = i(2z)^{-1} \sqrt{\frac{1}{2}\pi - z^{-1}}$$

which becomes arbitrarily small if $|z|$ increases arbitrarily. That is, if the desired accuracy is not too great one may restrict oneself to the first approximation of the zeros.

In consequence of (2.9) the zeros of the equation

$$(5.8) \quad \cos z = S\sqrt{z}$$

are the first approximation of the zeros $z_n^* = x_n^* + iy_n^*$ of $S(z)$. In a manner similar to the preceding one we obtain from (5.8)

$$(5.9) \quad y_{0,n}^* = \log(2\pi\sqrt{n}), \quad n = 1, 2, \dots$$

and

$$(5.10) \quad x_{0,n}^* = 2n\pi - \frac{\log(2\pi\sqrt{n})}{4n\pi}, \quad n = 1, 2, \dots$$

These formulas are also of a relatively high degree of accuracy. For the smallest zero of $S(z)$ the error is about 5 per cent for the imaginary part and 1 per cent for the real part. The error is much smaller for the larger zeros, e.g. about 1 per cent for the imaginary part and 0.3 per cent for the real part of the second zero, etc.

From (5.9) and (5.10) we find

$$(5.11) \quad \lim_{n \rightarrow \infty} y_n^*/x_n^* = 0.$$

6. More exact determination of the zeros. In the preceding section first approximations of the zeros of the Fresnel integrals $C(z)$ and $S(z)$ were obtained from (2.9). As follows from Theorem 3 and (2.10) the expansion (2.9) permits also a more exact determination of those zeros. In the case of the smallest zero (2.10) yields $m = 2$, i.e. the greatest possible accuracy is obtained if we take the constant term and the next 4 terms of (2.9) and determine the smallest zero of the function thus obtained. In the case of the second zero we have from (2.10) $m = 6$, i.e. we have to take the constant term and the next 12 terms of (2.9), etc. However, even the simplest of the equations which we obtain in this manner is too complicated and cannot be solved immediately. But there is another way which will turn out to be very simple.

Let us first consider the Fresnel integral $C(z)$. We start from the values obtained from (5.5) and (5.6) and improve those values by applying the Newton method. The values $z_{1,n}$ thus obtained from $z_{0,n}$ are more accurate approximations of the zeros of (5.1). We now apply the Newton method several times and, from step to step, we always take into account one more term of (2.9a). Let us denote by $f_p(z)$ the function which is obtained by taking the constant term and the next p terms of (2.9a). The zero of the equation $f_p(z) = 0$ which is contained in the strip S_n' (cf. §5) will be denoted by $z_{p,n}$. The derivative of $f_p(z)$ is given by the simple expression

$$(6.1) \quad f_p'(z) = z^{-1}(\cos z + h_p(z)), \quad p = 1, 2, \dots,$$

where $h_p(z)$ is of the form $k_1 z^{-p} \sin z$ or $k_2 z^{-p} \cos z$, k_1 and k_2 denoting certain constants; all other terms drop out in pairs; $h_p(z)$ is small in comparison with $\cos z$. If $z_{p-1,n}$ denotes the zero of the equation $f_{p-1}(z) = 0$ in S_n' then $f_p(z_{p-1,n})$ consists of one term only, namely, of the last term of (2.9a) under consideration. The function $\tan z_{p,n}$, occurring as a factor in some of the Newton quotients f/f' , may be replaced by i ; this simplification is the same as that in the preceding section where we omitted the functions indicated by Landau symbols.

In the case of the Fresnel integral $S(z)$ the reasoning is exactly the same as in the case of $C(z)$.

The procedure yields a finite sequence $z_{1,n}, z_{2,n}, \dots$ of approximative values of the zero z_n of $C(z)$. The terms of this sequence are recursively determined by the following simple relation:

$$(6.2) \quad z_{p+1,n} = z_{p,n} + c_p(2z_{p,n})^{-p}, \quad p = 1, 2, \dots,$$

where

$$\begin{aligned} c_{2q+1} &= (-1)^q 1.3 \dots (4q+1), & q &= 0, 1, \dots, \\ c_{2q} &= (-1)^{q-1} 1.3 \dots (4q-1)i, & q &= 1, 2, \dots \end{aligned}$$

For $S(z)$ we similarly find

$$(6.3) \quad z_{p+1,n}^* = z_{p,n}^* + c_p(2z_{p,n}^*)^{-p}, \quad p = 1, 2, \dots,$$

where the constants c_p are the same as in the preceding formula. Of course the numerical values of the different correction terms are entirely different in both cases, since $z_{p,n}$ differs from the corresponding approximate value $z_{p,n}^*$. For every fixed value of p the corresponding correction term decreases monotonely with increasing n . Since, for fixed n and p , $|z_{p,n}^*| > |z_{p,n}|$, the absolute value of the correction term $c_p(2z_{p,n}^*)^{-p}$ is smaller than that of $c_p(2z_{p,n})^{-p}$, but greater than that of $c_p(2z_{p,n+1})^{-p}$.

7. Further properties of the zeros. From the preceding results we can draw some conclusions which might be of interest. Let us compare the zeros of $C(z)$ with those of $S(z)$. From (5.5), (5.6), (5.9), and (5.10) we find that not only the sequences $(y_{0,n})$, $(y_{0,n}^*)$, (γ_n) , (γ_n^*) , where $\gamma_n^* = (4n\pi)^{-1} \log(2\pi\sqrt{n})$,

are monotone, but also the sequences $y_{0,1}$, $y_{0,1}^*$, $y_{0,2}$, $y_{0,2}^*$, \dots and γ_1 , γ_1^* , γ_2 , γ_2^* , \dots . We thus obtain

THEOREM 5. *The zeros of $C(z)$ and $S(z)$ are (asymptotically) located on one and the same logarithmic curve, in alternating order; this curve can be represented in the form*

$$(7.1) \quad y = \pm \frac{1}{2} \log 2\pi x.$$

In consequence of (5.3), (5.7) and (5.11), the series

$$(7.2) \quad \sum_{n=1}^{\infty} |z_n|^{-1}, \quad \sum_{n=1}^{\infty} |z_n^*|^{-1}$$

are minorants of the harmonic series. Hence the series

$$(7.3) \quad \sum_{n=1}^{\infty} |z_n|^{-1-\epsilon}, \quad \sum_{n=1}^{\infty} |z_n^*|^{-1-\epsilon}, \quad \epsilon > 0,$$

converge, but the series (7.2) diverge. The functions $z^{\frac{1}{2}}C(z)$ and $z^{\frac{1}{2}}S(z)$ thus are entire functions of first order of divergence class.

According to the order of $z^{\frac{1}{2}}C(z)$ and $z^{\frac{1}{2}}S(z)$ the exponent of the exponential function contained in the Weierstrass product of these functions can at most be a linear function of z . It can be proved that, in our case, this function is actually a constant. Since

$$\lim_{|z| \rightarrow 0} z^{-1/2} C(z) = 2, \quad \lim_{|z| \rightarrow 0} z^{-3/2} S(z) = \frac{2}{3},$$

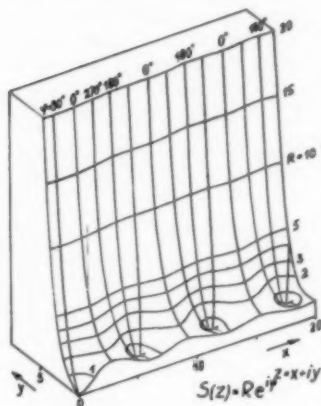
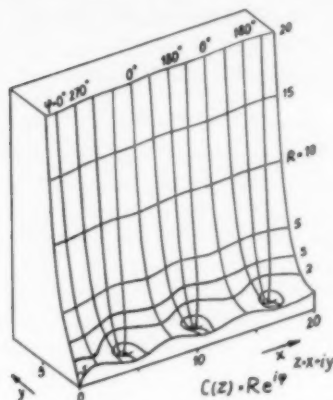
the Weierstrass products of the Fresnel integrals have the form

$$(7.4) \quad C(z) = 2z^{1/2} \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{z_n^2}\right) \left(1 - \frac{z^2}{z_n^{*2}}\right),$$

$$S(z) = \frac{2}{3} z^{3/2} \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{z_n^2}\right) \left(1 - \frac{z^2}{z_n^{*2}}\right).$$

8. Distribution of the function values of the Fresnel integrals. Tables of the function values of the Fresnel integrals for complex values of the argument have been communicated in (10; 11). Both functions have a similar behaviour which can be most simply described by characterising the geometric form of the surfaces $F(C)$ and $F(S)$ of the absolute value of $C(z)$ and $S(z)$.

In consequence of the maximum principle, the real extreme values of $C(z)$ and $S(z)$ correspond to saddle points of $F(C)$ and $F(S)$. Because of (2.4), both surfaces are symmetric with respect to the x -axis. Any zero is a singular point of the surfaces which, in a small neighbourhood of such a point, behave like a circular cone Z ; the angle between the generators of Z and the xy -plane is about $\frac{1}{2}\pi$, as can be seen from the Taylor series development of $C(z)$ and $S(z)$ at a zero under consideration. Any other point which corresponds to a finite value of z is a regular point of the surfaces. It can be seen from (2.9) that the surfaces ascend rapidly for large y . The tangent of the angle $\alpha(z)$



between the direction of maximum slope and the xy -plane is asymptotically given by the expression

$$(8.1) \quad \tan \alpha(z) = \frac{1}{2} |z|^{-1/2} e^{\pi/4}.$$

Along the lines $y = \text{const.}$, the maximal slope decreases with increasing x . There exist real points of inflection of the curves $C(x)$ and $S(x)$ at $x = n\pi - \delta(n)$ and $x = (2n + 1)\pi/2 - \delta^*(n)$, respectively, where $\delta(n)$ and $\delta^*(n)$ are positive and monotonely decreasing functions of n . All these points are isolated parabolic points of the surfaces $F(C)$ and $F(S)$. The surfaces consist of domains of positive and negative Gaussian curvature which are bounded and separated from each other by curves whose points are parabolic ("parabolic curves"). These curves can be obtained from

$$(8.2) \quad \Re(f''^2/f''f) - 1 = 0,$$

(cf. 21), where $f = C(z)$ or $S(z)$, respectively. Through any zero there passes exactly one of those curves; the curves remain always in a neighbourhood of the curves of constant phase $\pi/2$ and $3\pi/2$ with which they asymptotically coincide.

9. Tables. The methods developed in this paper enable us to investigate and to calculate the zeros of the Fresnel integrals in a simple manner.

We did not consider the functions $c(z)$ and $s(z)$ defined by (2.7). Although these functions are very simply related to the Fresnel integrals, their behaviour is different from that of $C(z)$ and $S(z)$. Since $c(z)$ and $s(z)$ also occur in connection with many practical problems they should eventually be studied more in detail; this will be done in a subsequent paper.

Zeros $z_n = x_n \pm iy_n$ of $C(z)$

n	x_n	y_n	n	x_n	y_n	n	x_n	y_n	n	x_n	y_n	n	x_n	y_n
1	4.62	1.68	11	67.53	3.03	21	130.36	3.35	31	193.20	3.55	41	256.03	3.69
2	10.94	2.11	12	73.81	3.07	22	136.65	3.38	32	199.48	3.57	42	262.32	3.70
3	17.24	2.34	13	80.10	3.11	23	142.93	3.40	33	205.77	3.58	43	268.60	3.72
4	23.53	2.50	14	86.38	3.15	24	149.21	3.42	34	212.05	3.60	44	274.88	3.73
5	29.82	2.62	15	92.66	3.18	25	155.50	3.44	35	218.33	3.61	45	281.17	3.74
6	36.11	2.71	16	98.95	3.22	26	161.78	3.46	36	224.62	3.63	46	287.45	3.75
7	42.40	2.79	17	105.23	3.25	27	168.06	3.48	37	230.90	3.64	47	293.73	3.76
8	48.68	2.86	18	111.51	3.28	28	174.35	3.50	38	237.18	3.65	48	300.02	3.77
9	54.96	2.92	19	117.80	3.30	29	180.63	3.52	39	243.47	3.67	49	306.30	3.78
10	61.25	2.98	20	124.08	3.33	30	186.92	3.53	40	249.75	3.68	50	312.58	3.79

Zeros $z_n^* = x_n^* + iy_n^*$ of $S(z)$

n	x_n^*	y_n^*	n	x_n^*	y_n^*	n	x_n^*	y_n^*	n	x_n^*	y_n^*	n	x_n^*	y_n^*
1	6.20	1.74	11	69.09	3.04	21	131.93	3.36	31	194.77	3.55	41	257.60	3.69
2	12.51	2.16	12	75.38	3.08	22	138.22	3.38	32	201.05	3.57	42	263.89	3.71
3	18.81	2.37	13	81.66	3.12	23	144.50	3.41	33	207.34	3.59	43	270.17	3.72
4	25.10	2.52	14	87.95	3.16	24	150.79	3.43	34	213.62	3.60	44	276.45	3.73
5	31.38	2.63	15	94.23	3.19	25	157.07	3.45	35	219.90	3.62	45	282.74	3.74
6	37.67	2.72	16	100.52	3.22	26	163.35	3.47	36	226.19	3.63	46	289.02	3.75
7	43.95	2.80	17	106.80	3.25	27	169.64	3.49	37	232.47	3.64	47	295.30	3.76
8	50.24	2.87	18	113.08	3.28	28	175.92	3.50	38	238.75	3.66	48	301.59	3.77
9	56.52	2.93	19	119.37	3.31	29	182.20	3.52	39	245.04	3.67	49	307.87	3.78
10	62.81	2.99	20	125.65	3.34	30	188.49	3.54	40	251.32	3.68	50	314.15	3.79

REFERENCES

1. A. Bloch, *Sur les intégrales de Fresnel*, Darboux Bull. (2), 46 (1922), 34-35.
2. M. Brelot, *Quelques propriétés des fonctions de Gilbert et de la spirale de Cornu*, Bull. Sci. Math. (2), 61 (1937), 133-160.
3. W. Buerck und H. Lichte, *Untersuchungen über die Laufzeit in Vierpolen und die Verwendbarkeit der Gleitfrequenzmethode*, Elektr. Nachr. Technik, 15 (1938), 78-101.
4. M. F. Egan, *Gamma functions and Fresnel's integrals*, Math. Gazette, 19 (1935), 366-367.
5. F. Emde, *Kurvenlineale*, Instrum. Kunde, 58 (1938), 400-413.
6. B. de Haan, *Nouvelles tables d'intégrales définies* (Leyden, 1867).
7. A. L. Higgins, *The Transition Spiral and its Introduction to Railway Curves* (London, 1921).
8. P. Humbert, *Sur les intégrales de Fresnel*, Bull. Soc. Sc. Cluj, 7 (1934), 530-534.
9. V. Jamet, *Sur les intégrales de Fresnel*, Nouv. Ann. (4), 3 (1903), 357-359.
10. E. Kreyszig, *Ueber den allgemeinen Integralsinus $Si(z, \alpha)$* , Acta Math., 85 (1951), 117-181.
11. E. Kreyszig, *Der allgemeine Integralkosinus $Ci(z, \alpha)$* , Acta Math., 89 (1953), 107-131.
12. Th. Laible, *Höhenkarte des Fehlerintegrals*, Z. angew. Math. Physik, 2 (1951), 484-486.
13. D. N. Lehner, *Cornu's spiral as a transition curve*, Cal. Inst. Tech., 3 (1904), 71-82.
14. A. Lindstedt, *Zur Theorie der Fresnelschen Integrale*, Ann. Physik (2), 17 (1882), 720-725.
15. W. Magnus and F. Oberhettinger, *Formulas and Theorems for the Special Functions of Mathematical Physics* (New York, 1949).
16. N. Nielsen, *Theorie des Integrallogarithmus und verwandter Transzendenten* (Leipzig, 1907).
17. L. Oerley, *Übergangsbogen bei Strassenkrümmungen* (Berlin, 1937).
18. N. A. Oumoff, *Interprétation géométrique des intégrales de Fresnel*, J. de Phys. (3), 6 (1897), 281-289.
19. Th. Poeschl, *Das Anlaufen eines einfachen Schwingers*, Ing. Archiv, 4 (1933), 98-102.
20. H. Salinger, *Zur Theorie der Frequenzanalyse mittels Suchtons*, Elektr. Nachr. Technik, 6 (1929), 293-302.
21. E. Ullrich, *Betragsflächen mit ausgezeichnetem Krümmungsverhalten*, Math. Z., 54 (1951), 297-328.
22. J. Zbornik, *Asymptotische Entwicklungen für Fresnelsche Integrale und verwandte Funktionen und ihre Anwendungsmöglichkeiten bei der Berechnung spezieller Raketenbahnen*, Z. angew. Math. Physik, 5 (1954), 345-351.

University of Ottawa and
Ohio State University

A MINIMUM-MAXIMUM PROBLEM FOR DIFFERENTIAL EXPRESSIONS

D. S. CARTER

1. Introduction. In the study of approximate methods for solving ordinary differential equations, an interesting question arises. To state it roughly for a single first order expression, let $y_0(t)$ be the solution of the equation

$$(1.1) \quad f(t, y, \dot{y}) = 0$$

which satisfies the initial condition $y(a) = \eta_a$. Let η_b be an *approximation* to the value of y_0 at a later time, $t = b$. Unless this approximation is exact, there is no continuous function which satisfies (1.1), together with the two boundary conditions

$$(1.2) \quad y(a) = \eta_a, \quad y(b) = \eta_b.$$

The question is whether there exists a continuous function satisfying (1.2), for which the maximum absolute value of $f(t, y, \dot{y})$ on the interval $[a, b]$ is minimized; and if so, how to find it.

Stated for higher order and multiple-component systems, this problem should find application in engineering and some branches of "operations analysis." For example, in a dynamical system being driven between two preassigned points in phase-space, it may be of interest to minimize the peak value of a stress which is expressible, through accelerations or friction, as the absolute value of a differential form.

A moment's reflection provides some insight into the nature of the solution. If there is a solution $y^*(t)$, then $f(t, y^*, \dot{y}^*)$ must be constant in absolute value, so that $|f|$ is everywhere equal to its supremum. Otherwise it would be possible to decrease $|f|$ near its maxima at the expense of increases elsewhere. This does not mean that f itself need be constant. In fact, it turns out that the value of f at the solution will generally have a number of discontinuities of sign, especially for higher order expressions.

For the first order system (1.1), (1.2), this situation is illustrated by the following rough variational argument. Let $y(t)$ satisfy (1.2), and let the partial derivative $f_y(t, y, \dot{y})$ be constant in sign throughout $[a, b]$. The first variation of f corresponding to a variation δy is

$$\delta f = f_y \delta y + f_{\dot{y}} \delta \dot{y}.$$

Received January 6, 1956. Work performed under the auspices of the U.S. Atomic Energy Commission.

The author wishes to express his indebtedness to J. Lehner and S. Ulam of the Los Alamos Scientific Laboratory for helpful discussions of this problem.

Solving for δy , and using the fact that $\delta y = 0$ at $t = a, b$, we have

$$\int_a^b ds [\delta f(s)/f_s(s)] \exp \left\{ \int_b^s dr f_s(r)/f_s(r) \right\} = 0,$$

which is a necessary and sufficient condition for the admissibility of δf . Since neither f_s nor the exponential factor changes sign, this requires little more than that δf shall change sign. Unless f is constant throughout $[a, b]$ it is easy to construct an admissible δf differing in sign everywhere from f , so $\sup|f + \delta f| < \sup|f|$. Conversely, if f is everywhere constant, then δf must agree in sign with f on part of the interval; and $\sup|f + \delta f| > \sup|f|$ for every non-vanishing δy . That is to say, $\sup|f|$ has at least a *local* minimum at $y = y^*$ if and only if y^* satisfies (1.2) and

$$f(t, y^*, \dot{y}^*) = c$$

for some constant c . Allowing for variation of c , this equation has a two-fold infinity of solutions; and in the simplest cases this is just enough to yield a unique solution satisfying (1.2).

The assumption that f_s be constant in sign may be relaxed simply by taking

$$f(t, y^*, \dot{y}^*) = \pm c,$$

where the sign agrees with that of f_s .

The object of this paper is to present a complete theory for linear systems alone, of which the chief results are contained in Theorem 1, §5, and Theorem 2, §6. Detailed consideration of non-linear cases is left for later publication. Meanwhile, rough extensions of the present theory by variational arguments (i.e., linearization near the solution) will undoubtedly yield correct results for most applications.

2. Formulation of the first order problem. All functions encountered are real, single-valued, and defined on a closed, bounded interval $I = [a, b]$.

By a "vector" we mean an ordered n -tuple of functions, $f = (f^i)$, for some fixed n . Similarly, a "matrix" is an nxn square matrix of functions. A vector or matrix is said to be continuous, or summable, etc., if and only if each of its components has that property. For matrix operations, vectors are regarded as rows or columns according to context.

Two spaces are fundamental to the discussion. The trial solutions $y(t)$ are taken from the space Y of absolutely continuous vectors; the values of the differential expressions lie in the space M of measurable vectors. Elements of M are regarded as equivalent, written $f \sim g$, when they are equal almost everywhere.

For all vectors we define the "vector absolute value"

$$(2.1) \quad |f| = |(f^i)| = (|f^i|),$$

and for elements of M we have the "maximum-norm"

$$(2.2) \quad \|f\| = \sup \{ \text{ess. sup. } |f^i| \mid i = 1, \dots, n \},$$

where the essential supremum of each component is taken over I .

Given an ordered n -tuple $J = (J^i)$ of subsets of I , the vector whose components are the characteristic functions of the corresponding sets J^i is called the characteristic vector of J .

To define the problem we are given

(a) a first order differential operator

$$(2.3) \quad \mathfrak{L}y = A[\dot{y} + By + c]$$

which serves to map Y into M . A is an almost everywhere finite and non-singular matrix, whose inverse A^{-1} , together with the matrix B and vector c , are Lebesgue summable.

(b) an ordered n -tuple J of measurable subsets of I , at least one of whose components has positive measure. The components of $\mathfrak{L}y$ will be free to vary on the corresponding component sets J^i , but must vanish almost everywhere on the complements $I - J^i$.

(c) a pair of initial conditions for vectors $y \in Y$, one for each endpoint of I :

$$(2.4) \quad y(a) = \eta_a, \quad y(b) = \eta_b.$$

Let $F \subset M$ be the space of equivalence classes of essentially bounded vectors f , whose components f^i vanish almost everywhere on the corresponding sets $I - J^i$. Let $X \subset Y$ be the set of all absolutely continuous vectors x which satisfy the initial conditions (2.4), and such that $\mathfrak{L}x \in F$. Then a vector x_0 is said to be a solution of the minimum-maximum problem if and only if $x_0 \in X$ and

$$(2.5) \quad \|x_0\| = \inf \{ \|x\| \mid x \in X \}.$$

It is important to notice that a given problem may be inconsistent, in the sense that the set X is empty. (For example, $n > 1$, $\mathfrak{L}y \equiv \dot{y}$. J^1 is nul, and $\eta_a^1 \neq \eta_b^1$. Here $\dot{x}^1 \sim 0$ for $x \in X$, which contradicts the requirement that $x^1(a) \neq x^1(b)$.) However, we will find that every consistent problem has a solution.

3. The set $G = \mathfrak{L}X$. Although the solutions are defined in terms of X , it is more convenient to work with the image G of X under \mathfrak{L} . This is possible because of

LEMMA 1. *The sets X and G are in 1:1 correspondence, for \mathfrak{L} has a unique inverse on G .*

The proof consists of the observation that for every $f \in F$ the differential equation $\mathfrak{L}y \sim f$ has a unique solution satisfying either of the two initial conditions (2.4).¹

¹The properties of solutions of all the differential equations considered in this section are essentially the same as if the coefficients were continuous. They may be derived by a slight extension of the standard methods (1; 2, chap. IX).

In view of this correspondence, the solutions may be defined in terms of G : g_0 is a solution if and only if $g_0 \in G$ and

$$\|g_0\| = \inf \{\|g\| \mid g \in G\}.$$

The remainder of this section is devoted to a derivation of the necessary and sufficient conditions stated in Lemma 2 for an element of F to belong to G . First we will find expressions for the two inverses of \mathcal{L} corresponding to the two initial conditions at a and b . The required conditions follow from the fact that these inverses are equal on G .

Let W be the space of solutions of the homogeneous equation

$$\dot{y} + By \sim 0,$$

and let Z be the space of solutions of the adjoint equation

$$\dot{y} - yB \sim 0.$$

W and Z are both n -dimensional subspaces of Y . For every pair $w \in W$ and $z \in Z$ the "scalar product"

$$zw = \sum_{i=1}^n z^i(t)w^i(t)$$

is constant, independent of t .

If n vectors $\{e_j\}$ form a basis for W , they may be combined into a matrix E which is everywhere non-singular. The matrix

$$K(t, s) = E(t) E^{-1}(s)$$

is independent of the choice of basis, and plays the role of "translation operator" in W and Z :

$$(3.1) \quad w(t) = K(t, s) w(s), \quad z(s) = z(t) K(t, s).$$

For every $f \in F$, the pair y_a, y_b of solutions of

$$\mathcal{L}y \sim f$$

which satisfy the initial conditions (2.4) at a and b , respectively, are given by

$$(3.2a) \quad y_\alpha(t) = w_\alpha(t) + \int_a^t K(t, s) [A^{-1}(s) f(s) - c(s)] ds, \quad \alpha = a, b,$$

where $w_\alpha(t)$ is the element of W which is equal to η_α at $t = \alpha$:

$$(3.2b) \quad w_\alpha(t) = K(t, \alpha) \eta_\alpha.$$

Now $y_a = y_b$ if and only if $f \in G$, so we have, with the help of (3.1):

$$(3.3)^2 \quad \text{LEMMA 2. } g \in G \text{ if and only if } g \in F \text{ and} \\ u(t) = \int K(t, s) A^{-1}(s) g(s) ds$$

²Integrals written without limits mean integrals from a to b .

where u is an element of W which is determined by the vector c and the initial values η_a and η_b :

$$(3.4) \quad u(t) = w_b(t) - w_a(t) + \int K(t, s) c(s) ds.$$

Moreover, (3.3) is equivalent to the condition

$$(3.5) \quad zu = \int \xi(s) g(s) ds \quad \text{for every } z \in Z$$

where, for each $z \in Z$, the vector ξ is defined by

$$(3.6) \quad \xi = zA^{-1}.$$

4. The function $\mu(z)$. Our plan is to reduce the problem so that its solution amounts to maximizing a function on the n -dimensional space Z , rather than minimizing a function on the infinite-dimensional space F . This reduction depends on the following inequality, which is a direct consequence of (3.5) and the definition of F .

Let j be the characteristic vector of J . Then for every pair $z \in Z$ and $g \in G$,

$$(4.1) \quad |zu| \leq \|g\| \int |\xi| j ds.$$

Now consider the set

$$(4.2) \quad V = \{z | z \in Z \text{ and } \int |\xi| j ds = 0\},$$

which is clearly a linear subspace of Z . When its dimension is less than n , the function

$$(4.3) \quad \mu(z) = zu / \int |\xi| j ds$$

is defined on the complement CV of V in Z , and satisfies the inequality

$$(4.4) \quad \sup \{|\mu(z)| | z \in CV\} \leq \inf \{\|g\| | g \in G\}.$$

We will see that equality actually holds, as long as μ is bounded.

Some useful properties of $\mu(z)$ are contained in the proof of

LEMMA 3. When $V \neq Z$, $\mu(z)$ is bounded on CV if and only if

$$(4.5) \quad zu = 0, \text{ for every } z \in V.$$

And if μ is bounded, then $|\mu|$ attains its supremum on CV .

Proof. Choose components with respect to an arbitrary basis as coordinates in Z , so that Z becomes homeomorphic to Euclidean n -space. Then both numerator and denominator in the expression (4.3) are clearly continuous on Z . If $\mu(z)$ is bounded on CV , then since V forms the boundary of CV , the numerator must vanish with the denominator as $z \rightarrow V$, and (4.5) must be true.

Conversely, let (4.5) be satisfied. Let U be any linear subspace of Z complementary to V , so that every $z \in CV$ is expressible in the form $z = z_1 + z_2$, where $z_1 \in U$, $z_2 \in V$, and $z_1 \neq 0$. Substituting into (4.3), we find that $\mu(z_1 + z_2) = \mu(z_1)$. Moreover, for every real α , $\mu(\alpha z) = \pm \mu(z)$, where the

sign agrees with that of α . Hence μ assumes all its functional values for z on the unit sphere in U . The proof is completed by recalling that a continuous function on a closed and bounded set in Euclidean space is bounded, and attains its supremum on that set.

5. Solution of the first order problem. The way is now prepared for

THEOREM 1. *The problem is consistent if and only if the condition (4.5) of Lemma 3 is satisfied. Moreover, every consistent problem has a solution as follows:*

- (a) *in the trivial case $u = 0$, there is a unique solution $g_0 \sim 0$.*
- (b) *in the "degenerate case" $V = Z$, the problem is consistent only if $u = 0$.*
- (c) *in all other cases the function $|\mu(z)|$ has a positive supremum which is attained for some $z \in CV$. Let z_0 be any such point, and let*

$$(5.1) \quad \mu_0 = \mu(z_0), \quad \xi_0 = z_0 A^{-1}.$$

Then every solution has the property

$$(5.2) \quad ||g_0|| = |\mu_0|$$

and its components are uniquely determined on the sets

$$(5.3) \quad J_0^i = \{t | t \in J^i, \quad \xi_0^i(t) \neq 0\}$$

by the equivalences

$$(5.4) \quad g_0^i \sim \mu_0 \xi_0^i / |\xi_0^i|.$$

Thus each g_0^i is equal in absolute value to the constant $|\mu_0|$ almost everywhere on J_0^i ; and the solution is unique if every J_0^i has the same measure as the corresponding J^i .

Proof. The problem is inconsistent if (4.5) is violated, for otherwise the inequality (4.1) provides an obvious contradiction. Now the result for case (b) follows, since (4.5) implies $u = 0$ when $V = Z$. And in view of Lemma 2, the result for case (a) is trivial.

The rest of the proof deals with case (c), in which $u \neq 0$ and $V \neq Z$. Assuming that (4.5) is satisfied, we will construct a $g_0 \in F$ that satisfies (3.3), (5.2), and (5.4). Then in view of Lemma 2, $g_0 \in G$, so (4.5) implies consistency. And by virtue of (4.4), g_0 is also a solution. To show that the components of every solution satisfy (5.4), let g_1 be any other solution. Since

$$|g_1^i| \leq |g_0^i| = |\mu_0|$$

almost everywhere on J_0^i , there exists a vector e such that $g_0^i - g_1^i \sim e_0^i g_0^i$ on each J_0^i , where $0 \leq e^i \leq 2$. Since g_0 and g_1 both satisfy (3.5), it follows that

$$\int \xi_0 (g_0 - g_1) ds = \mu_0 \sum_i \int_{J_0^i} |\xi_0^i| e^i ds = 0.$$

Hence each $e^i \sim 0$, and $g_1^i \sim g_0^i$ on J_0^i .

To construct g_0 , consider the vectors $h \in F$ and $v \in W$ defined by

$$h^i = \begin{cases} \mu_0 \xi_0^i / |\xi_0^i| & \text{on } J_0^i, \\ 0 & \text{elsewhere,} \end{cases}$$

$$v = \int K(t, s) A^{-1}(s) h(s) ds$$

If $v = u$, take $g_0 \sim h$. If $v \neq u$, define a new problem with the same spaces W and Z but with u and J replaced by

$$u' = u - v,$$

$$J'^i = J^i - J_0^i.$$

It is shown below that the inequality

$$(5.5) \quad |zu'| \leq |\mu_0| \int |\xi| j' ds$$

holds for every $z \in Z$, where j' is the characteristic vector of J' . Then it is easy to check that the new problem falls under case (c), and that

$$\sup |\mu'(z)| \leq |\mu_0|.$$

Moreover, the dimension of the new space V' exceeds that of V , for $V \subset V'$ and $z_0 \in V'$, while $z_0 \notin V$.

The procedure just outlined may be repeated to give a sequence of problems, which terminates as soon as the condition $v = u$ is satisfied. Termination occurs after at most $n - d$ steps, where d is the dimension of the original V . For if the sequence continues until V is of dimension $n - 1$, then $V' = Z$, and it follows from (5.5) that $u' = u - v = 0$.

If h_j denotes the vector h for the j th step, it is easy to check that

$$(5.6) \quad g_0 \sim \sum_j h_j$$

has all the required properties.

To prove (5.5), note first that

$$(5.7) \quad (z_0 + \alpha z) u / \mu_0 \leq \int |\xi_0 + \alpha \xi| j ds$$

for every z and every real α . Keeping z fixed, consider the vector k and the sets L_α defined by

$$k^i = \begin{cases} \xi_0^i \xi_0^i / |\xi_0^i| & \text{on } J_0^i, \\ 0 & \text{elsewhere,} \end{cases}$$

$$L_\alpha^i = \{t | t \in J_0^i \text{ and } |\xi_0^i| + \alpha k^i < 0\}.$$

Then on J_0^i

$$|\xi_0^i + \alpha \xi^i| = |\xi_0^i / |\xi_0^i| (|\xi_0^i| + \alpha \xi_0^i \xi_0^i / |\xi_0^i|)| = ||\xi_0^i| + \alpha k^i|$$

and on J'^i ,

$$|\xi_0^i + \alpha \xi^i| = |\alpha| |\xi^i|.$$

Hence, if j_0 and l_α are the characteristic vectors of J_0 and L_α respectively,

$$(5.8) \quad \int |\xi_0 + \alpha \xi| j ds = \int [|\xi_0| + \alpha k] j_0 ds + |\alpha| \int |\xi| j' ds - 2 \int [|\xi_0| + \alpha k] l_\alpha ds.$$

Now in view of (3.1) and (3.6),

$$zv = \int \zeta(s) h(s) ds$$

and, since $\zeta h = \mu_0 k j_0$,

$$\int k j_0 ds = zv / \mu_0$$

Moreover,

$$\int |\zeta_0| j_0 ds = \int |\zeta_0| j ds = z_0 u / \mu_0.$$

Combining the last two equations with (5.8) and (5.7) and using the fact that

$$0 < -[|\zeta_0^t| + \alpha k^t] < |\alpha k^t| \quad \text{on } L_{\alpha}^t,$$

it follows that

$$\pm zu' / \mu_0 < \int |\zeta| j' ds + 2 \int |k| l_{\alpha} ds,$$

where the sign agrees with that of α .

For every monotone vanishing sequence $\{\alpha_j\}$, the corresponding sequences $\{L_{\alpha_j}^t\}$ are non-decreasing, and the infinite products $\prod_j L_{\alpha_j}^t$ contain only points at which the k^t are infinite. Hence the measure of each L_{α}^t vanishes with α . The proof is completed by taking the two limits $\alpha \rightarrow \pm 0$ in the last inequality.

6. Higher order systems. The theory is easily extended to cases in which the operator \mathcal{Q} is of higher order, by rewriting the given system in an equivalent first order form. The case of a single higher order expression provides an interesting illustration, in view of the special results obtained below.

Taking $n = 1$, let \mathcal{Q} be replaced by the m th order operator

$$(6.1) \quad \mathcal{Q}_m y = a \left(\frac{d^m y}{dt^m} + \sum_{j=m-1}^0 b_j \frac{d^j y}{dt^j} + c \right)$$

where a is almost everywhere finite, and $1/a$, the b_j , and c are summable functions. $\|\mathcal{Q}_m x\|$ is to be minimized over all functions x such that

(a) x and its first $m - 1$ derivatives are absolutely continuous, and assume given values at the endpoints of I .

(b) $\mathcal{Q}_m x$ is essentially bounded, and vanishes almost everywhere on $I - J_m$, where J_m is a given subset of positive measure.

On taking

$$y^i = \frac{d^{i-1} y}{dt^{i-1}}, \quad i = 1, \dots, m$$

this reduces to a first order problem in which $n = m$, all the J^i are empty except $J^m = J_m$, and the matrix A is diagonal, with unit diagonal elements except $a^{mm} = a$. Accordingly, the function μ has the form

$$\mu(z) = zu / \int_{J_m} |z^m / a| ds$$

where the integral extends over J_m .

It is shown below that for every $z \neq 0$ the component z^m has no more than a finite number of zeros on I . Hence the set V consists of the single element $z = 0$, and the condition (4.5) is always satisfied. And if $|\mu|$ attains its supremum at z_0 , the set

$$J_0^m = \{t | t \in J_m, z^m/a \neq 0\}$$

differs from J_m by a set of measure zero. Thus we have

THEOREM 2. *For a single differential expression of the form (6.1) the problem is always consistent and has the unique solution*

$$g_{m0} \sim \begin{cases} z_0^m a / |z_0^m a| & \text{on } J_m, \\ 0 & \text{on } I - J_m. \end{cases}$$

And except in the trivial case $u = 0$, the solution has a finite or infinite number of sign changes according as the factor $a(t)$ changes sign a finite or infinite number of times on J_m .

Proof. Let z be any element of Z for which z^m has an infinity of zeros on I . These zeros must have a limit point $\tau \in I$. We will show that if, for any j in the range $m \leq j < 1$, τ is a common limit point of the zeros of each component z^m, z^{m-1}, \dots, z^j , then τ is also a limit point of the zeros of z^{j-1} . Hence the zeros of every component have τ as a common limit point. Since z is continuous, it follows that every component vanishes at $t = \tau$, and hence that z vanishes identically.

Now the adjoint equations

$$z^i - b_{i-1} z^m + z^{i-1} \sim 0, \quad j \leq i \leq m,$$

may be regarded as equations for (z^j, \dots, z^m) with the inhomogeneous term $(z^{j-1}, 0, \dots, 0)$. If z^j, \dots, z^m all vanish at $t = \tau$, it follows from the analog of (3.2) that

$$z^j(t) = - \int_{\tau}^t k_j(t, s) z^{j-1}(s) ds$$

where k_j is continuous in both its arguments, and $k_j(\tau, \tau) = 1$. Hence, for sufficiently small $|t - \tau|$, the continuous function z^{j-1} must change sign between τ and each zero of z^j . This completes the proof.

REFERENCES

1. G. D. Birkhoff and R. E. Langer, *The boundary problems and developments associated with a system of ordinary differential equations of the first order*, Proc. Amer. Acad. Arts Sci. 58 (1923), 345-424.
2. E. J. McShane, *Integration* (Princeton, 1944).

Los Alamos Scientific Laboratory

A MIXED PROBLEM FOR NORMAL HYPERBOLIC LINEAR PARTIAL DIFFERENTIAL EQUATIONS OF SECOND ORDER

G. F. D. DUFF

In the theory of hyperbolic differential equations a mixed boundary value problem involves two types of auxiliary conditions which may be described as initial and boundary conditions respectively. The problem of Cauchy, in which only initial conditions are present, has been studied in great detail, starting with the early work of Riemann and Volterra, and the well-known monograph of Hadamard (4). A modern treatment of great generality has been given by Leray (7). In contrast mixed problems have received comparatively little attention, and the nature of the boundary conditions to be imposed on equations of order higher than the second is known only for equations in two independent variables (8). For second order normal hyperbolic equations both linear and non-linear, the problem has been studied, using the method of analytic approximation, by Schauder and Krzyzanski (5) who assigned as boundary condition that the unknown function should take given values on a timelike boundary surface. The monograph of Ladyshenskaya (6) treats certain cases of the problem where the normal derivative is given, for instance when the coefficients are independent of the time variable.

In this paper a different boundary condition is considered; this condition involves the derivative of the dependent variable in a given direction, which is defined on the boundary but is not tangential to the boundary. There are restrictions in the large on this direction, made necessary by the properties of certain families of characteristic surfaces. However, the condition includes as a special case the problem of the normal derivative, which arises in the theory of supersonic flow.

As in (5) the analytic case is treated first, by means of dominant power series. The nature of the boundary conditions is taken into account by a certain order of choice among the dominating series. For the non-analytic case a suitable modification of the estimates of (5) is arranged, while the construction of the solution is as before.

1. The mixed problem. We study the linear normal hyperbolic partial differential equation

$$(1.1) \quad L(u) = a^{ik} \frac{\partial^2 u}{\partial x^i \partial x^k} + b^i \frac{\partial u}{\partial x^i} + cu = f,$$

Received February 27, 1956.

The author is indebted to Professor J. Leray for an interesting and valuable discussion of this problem. He also acknowledges with thanks the helpful advice and criticism of Professor A. Robinson.

with one dependent variable u and N independent variables x^i ($i = 1, \dots, N$). Summation over repeated indices is understood in (1.1). The coefficients a^{ik} , b^i , c , and f are functions of the x^i , differentiable k times throughout the domain of x^i space to be considered. The normal hyperbolic character of (1.1) is expressed by the signature of the quadratic form

$$(1.2) \quad a^{ik} \xi_i \xi_k,$$

which signature is $(N - 1, 1)$ with one negative term.

With the Riemann metric

$$(1.3) \quad ds^2 = a_{ik} dx^i dx^k$$

based on the associate covariant tensor a_{ik} , we have a classification of directions v^i as spacelike, null, or timelike, according as

$$v^2 = a_{ik} v^i v^k = a^{ik} v_i v_k$$

is positive, zero or negative. Also surfaces $S: \phi(x^i) = 0$ shall be spacelike, null or timelike according as

$$(1.4) \quad a_{ik} \frac{\partial \phi}{\partial x^i} \frac{\partial \phi}{\partial x^k}$$

is negative, zero, or positive. The normal n^i is defined by

$$(1.5) \quad n^i = a^{ik} \frac{\partial \phi}{\partial x^k}.$$

Let $S: \phi(x^i) = 0$ be an initial spacelike surface and let $T: \psi(x^i) = 0$ be a timelike surface intersecting S in a rim C of $N - 2$ dimensions. We shall suppose that S is bounded by C and that T is bounded "toward the past" by C , in a suitable orientation of "time." Let G be the characteristic surface, passing through the rim C , which lies in the region enclosed by S and T , that is, which bounds the domain of dependence D_S on S according to the theory of the Cauchy problem. We note that G is composed of characteristic strips tangent to the rim C .

On S we assign values of u and $\partial u / \partial n$; these are the usual Cauchy data, and they determine u in D_S . On T we assign a boundary condition "of the second kind" as follows. Let v be a vector field defined on T and subject to restrictions stated below. Then, if the directional derivative of u in the direction of v is denoted by $\partial u / \partial v$, we set

$$(1.6) \quad \frac{\partial u}{\partial v} = v^i \frac{\partial u}{\partial x^i} = f(x^i),$$

where $f(x^i)$ is a datum function given on T .

On the rim C this datum function $f(x^i)$ and its derivatives transverse to C in T shall be subject to certain conditions of compatibility with respect to the given differential equation and Cauchy data. These conditions were imposed to ensure that the derivatives of u up to a certain order shall be

continuous across G . The value for u on C shall be that assigned by the Cauchy data, and the first compatibility condition is that f shall be equal to the value of $\partial u / \partial v$ on C derived from the Cauchy data. The second compatibility condition brings in the differential equation since it postulates that $\partial f / \partial t$ equal $\partial^2 u / \partial v \partial t$, the latter being calculated on C from the Cauchy data and the differential equation. Here t denotes a suitable timelike variable. Likewise the k th condition of compatibility determines the $(k-1)$ st derivative of f as equal to the corresponding derivative of u , calculated from the Cauchy data and the differential equation, by successive differentiation with respect to t and substitution of values already found.

If the first k compatibility conditions hold, it is evident that u and its first k derivatives are continuous across G at the rim C . Now the method of construction of the solutions leads to their being continuous across G , at every point of G . To show that the transverse derivatives of u up to order k are also continuous across G , we note that through each point of G there passes a bicharacteristic ray issuing from the rim C . If a solution u is continuous and has continuous tangential derivatives on G , then its first transverse derivative u_n is continuous along the entire bicharacteristic if it is so at one point. The conditions necessary for this conclusion will be satisfied in our construction, and we infer that u_n is continuous across G . In succession the higher transverse derivatives, up to and including order k , are proved continuous across G .

The initial and boundary conditions determine u in a larger region. Construct the retrograde characteristic cone C^p with vertex P . If C^p , T , and S together bound a region, then P lies in this domain of dependence on S and T . Since the Cauchy initial value problem can be regarded as solved, we subtract its solution from the dependent variable and so find a reduced boundary value problem which may be stated as follows. We must find a solution of (1.1), with $f = 0$, which vanishes on G_0 , satisfies (1.6) on T , and is defined in the region V intermediate to S and T . The compatibility condition is now that $f(x^i) = 0$ on C , while the corresponding condition of order k is that the derivative of $f(x^i)$ of order $k-1$ in a direction within T but transverse to C should vanish. We note that the rim C , being a subspace of S , is spacelike, and that all directions tangent to G are either spacelike or null. Indeed G , being an integral surface of a first order partial differential equation, is composed of the bicharacteristic curves passing through C which determine characteristic strips tangent to C .

Now let C_i be a family of spacelike $(N-2)$ -dimensional surfaces filling T , and such that $C_0 = C$. We may construct characteristic surfaces G_i containing C_i and these will fill up the region V . The condition which we impose on the vector field v is that v should not be tangent to the local G_i at any point of T .

That some restriction of the vector field relative to characteristic surfaces is necessary can be seen from the equation of the vibrating string:

$$u_{xx} = u_{tt}.$$

If we require $u(x, 0) = u_t(x, 0) = 0$ for $t = 0, x > 0$, and

$$\alpha(t)u_x + \beta(t)u_t = f(t), \quad x = 0, t > 0,$$

then

$$u(x, t) = \int_0^{t-x} \frac{f(\tau)}{\beta(\tau) - \alpha(\tau)} d\tau, \quad t > X,$$

and the denominator of the integrand vanishes if $\alpha(\tau) = \beta(\tau)$; that is, if the vector field takes the direction of the forward characteristic entering the region through T . This condition holds for all such equations in two variables; it is easily seen that the directional derivative along any forward characteristic is determined by the Cauchy problem with data taken at the instant t where the characteristic curve meets T .

If $N > 2$, the situation is more involved. For the analogous condition, namely that v should not touch G_t , we may ask: what conditions on v enable us to construct a family of spacelike varieties C_t on T such that v will never be tangent to the G_t constructed on C_t as base? Such a field will be called admissible.

To answer the question we recall that the G_t , being integral surfaces of a first order partial differential equation, are constructed as envelopes of portions of the characteristic conoids with vertex on C_t . The portion concerned is that part C_t^P of the forward half-cone lying in the interior of our region.

Let us consider the tangent space at a typical point P of T : $\psi \equiv X = X^{N-1} = 0$. In the surface T the tangent plane to C_t will take the form (with P as origin)

$$t - \sum c_\alpha x_\alpha = 0, \quad \alpha = 1, \dots, N-2,$$

where $\sum c_\alpha^2 < 1$ since C_t is spacelike. The tangent plane to the cone C_t^P in the full space, which also meets T tangent to C_t , will have the equation

$$t - \sum c_\alpha x_\alpha - \sqrt{(1 - \sum c_\alpha^2)} x = 0.$$

We must determine those regions of space such that there exist values of the c_α which render the function

$$f \equiv t - \sum c_\alpha x_\alpha - \sqrt{(1 - \sum c_\alpha^2)} x$$

consistently of a given sign.

There will later appear the restriction that the vector v should not touch T , and we shall, for convenience, make this assumption here. The region of space to be considered may now be taken as the side $x > 0$ of T . Two cases arise, according as

$$f_0[v] = v_t - v_x,$$

the initial value of f with the components of v substituted for the coordinates, is positive or negative. These correspond to v lying initially "later" than G or "earlier," and will be referred to as the positive and negative cases respectively.

Taking first the case of negative values, we shall minimize f with respect to the c_a . Since

$$\frac{\partial f}{\partial c_a} = -x_a + \frac{c_a}{\sqrt{(1 - \sum c_a^2)}} x$$

and

$$\frac{\partial^2 f}{\partial c_a \partial c_\beta} = \frac{\delta_{a\beta} x}{\sqrt{(1 - \sum c_a^2)}} + \frac{c_a c_\beta x}{(1 - \sum c_a^2)^{3/2}},$$

we find, first, that an extremum is present for

$$c_a = \frac{\pm x_a}{\sqrt{(x^2 + \sum x_a^2)}},$$

and secondly, that this is a minimum value. The actual minimum is therefore

$$f_{\min} = t - \sqrt{(x^2 + \sum x_a^2)}$$

which will be negative if

$$t < \sqrt{(x^2 + \sum x_a^2)}.$$

That is, points on the forward cone C^P_t , or within it, are excluded.

The positive case is a little different, since no true maximum of f exists. As we have $x > 0$ the third term of f is negative, and it follows that if

$$f_1 \equiv t - \sum c_a x_a$$

takes positive values for some c_a , then so does f for sufficiently small positive x . Now the sum in f will take its greatest value when $\sum c_a^2$ is allowed its greatest value. Thus we may take $\sum c_a^2 = 1$ and so find the extrema of

$$f_2 = t - \sum c_a x_a + \lambda(1 - \sum c_a^2).$$

Hence

$$\frac{\partial f_2}{\partial c_a} = -x_a - 2\lambda c_a = 0,$$

and so, with $\sum c_a^2 = \sum x_a^2 / 4\lambda^2 = 1$, we find

$$c_a = \frac{-x_a}{\sqrt{(\sum x_a^2)}}$$

and the maximum of f_1 is

$$f_{1\max} = t - \sum c_a x_a = t + \sqrt{(\sum x_a^2)}.$$

Thus we get positive values for f_1 , and since $\sum c_a^2 = 1$, also for f , provided

$$t > -\sqrt{(\sum x_a^2)}.$$

The bounding surface so defined is cylindrical in the x direction, and touches the cone C^P_t along its intersection with the surface T .

If the condition of not touching holds at a point P , then by continuity it holds in a neighbourhood of P . Over a compact portion of T we can find

uniform moduli of continuity, provided that the limiting cases mentioned above do not arise. A neighbourhood of uniform size can thus be defined, and the construction extended to the whole of the compact region by repeated application of the existence theorem.

This result may be summarized as follows:

LEMMA I. *A vector field v not tangent to G_0 or to T is admissible if*

(a) *being initially positive, it satisfies*

$$v_t > -\sqrt{(\sum v_a^2)} \quad \text{on } T;$$

(b) *being initially negative, it satisfies*

$$v_t < \sqrt{(v_x^2 + \sum v_a^2)} \quad \text{on } T.$$

We remark that the normal vector field, with one non-vanishing component $v_x > 0$, falls under case (b).

2. The analytic case. Let all coefficients in the differential equation, and the surfaces S and T , be analytic. Then characteristic surfaces such as the G_i are also analytic provided that the rims C_i are analytic. This can be arranged and will be assumed.

Before reducing the differential equation to a standard form (4, p. 76) we shall simplify the boundary condition

$$(2.1) \quad \frac{\partial u}{\partial v} = f.$$

Here f vanishes to order $k+1$ on C_0 according to the compatibility conditions. We note that the vector field u is not parallel to G_0 on T and thus we can construct a C^k function u_1 which vanishes on G_0 and also satisfies (2.1) on T . Subtracting this function from u , we obtain for the new dependent variable a differential equation of the form

$$(2.2) \quad L(u) = f_1$$

while the new boundary conditions are (cf. §6),

$$(2.3) \quad u = 0, \quad \text{on } G_0,$$

with

$$(2.4) \quad \frac{\partial u}{\partial v} = 0, \quad \text{on } T.$$

We now change the independent variables so as to give T the equation $x = x^{N-1} = 0$ while the analytic family of characteristic surfaces G_i have equation

$$G_i: t = x^N = \text{const.}$$

This forces the coefficient a^{NN} to vanish identically in the new system. Since the rim C_i is spacelike, and so never tangent to a characteristic direction,

we can choose the remaining variables x^1, \dots, x^{N-2} so that the bicharacteristics on G , are

$$(2.5) \quad x^p = \text{const.}, \quad p = 1, \dots, N-2.$$

This results in the vanishing of the coefficients a^{Np} . Following Hadamard (4), we divide by $a^{N, N-1}$ which cannot now vanish since $L(u)$ is not parabolic; and we replace u by

$$u \exp \left[\int b^N dx \right]$$

which causes the term in $L(u)$ containing $\partial u / \partial t$ to disappear. Then the differential equation becomes

$$(2.6) \quad \frac{\partial^2 u}{\partial x \partial t} = L_1(u) + f_2,$$

where the operator $L_1(u)$ contains no differentiations with respect to t . With this form of the equation Hadamard and others studied the indeterminacy of Cauchy's problem for characteristic surfaces.

The boundary conditions to go with (2.6) are now

$$(2.7) \quad u = 0 \text{ for } t = 0$$

and

$$(2.8) \quad \frac{\partial u}{\partial v} = v^i \frac{\partial u}{\partial x^i} = b^N u, \quad x = 0.$$

In order to express this latter condition more conveniently, we note that by hypothesis the component v^N does not vanish—this is our condition on the vector v . Dividing by v^N and transposing some terms, we have

$$(2.9) \quad \frac{\partial u}{\partial t} = \sum_{k=1}^{N-1} \beta^k \frac{\partial u}{\partial x^k} + hu, \quad x = 0.$$

We now expand u in a series of powers of t , and determine the coefficients in succession. Let

$$(2.10) \quad u = \sum_{n=0}^{\infty} u_n t^n, \quad f_2 = \sum_{n=0}^{\infty} f_n t^n,$$

and also let

$$L_1(u) = \sum_{n=0}^{\infty} t^n L_{n1}(u).$$

Then the u_n satisfy

$$(2.11) \quad n \frac{\partial u_n}{\partial x} = L_{01}(u_{n-1}) + f_{n-1} + \dots,$$

where the terms omitted contain the u_k ($k = 0, 1, \dots, n-2$). We have taken $n_0 \equiv 0$ to satisfy (2.7). Substituting these expansions into (2.8), we get the conditions

$$(2.12) \quad nu_n = \sum_k \beta_0^k \frac{\partial u_{n-1}}{\partial x^k} + h_0 u_{n-1} + \dots, \quad x = 0.$$

Here β_0^k and h_0 are initial terms in the expansions of the β^k and h in powers of t , while the terms omitted in (2.12) again contain u_k ($k = 0, 1, \dots, n-2$). Thus the u_n are uniquely determined by integration of (2.11) for successive values of n , in the form

$$(2.13) \quad nu_n(x) = nu_n(0) + \int_0^x [L_{01}(u_{n-1}) + f_{n-1} + \dots] dx',$$

and the functions so found are analytic in x as well as in the remaining variables.

The techniques of dominating series will now be applied to show that the series solution thus found is convergent in a certain domain. We note that the operations in (2.13) are such as to preserve any dominant relation; thus if we dominate the coefficients in (2.6) and (2.9) the new solution will dominate that already found. Now the two auxiliary conditions will be dominated in the following way. We shall seek a solution with positive coefficients of the dominating differential equation. This solution will automatically dominate the condition (2.7). We will also show that if the left side of (2.9) is computed (in the dominant case) it will dominate the right side, and therefore will dominate the actual condition (2.9). This requires a certain order of choice among the various dominating constants which will appear. The proof will also show that the series has a radius of convergence independent of the function f_2 in (2.6), and hence independent of the data prescribed for the original problem.

We choose as origin a point of C_0 and set

$$y = x^1 + \dots + x^{N-2},$$

and let ρ, σ be sufficiently small positive numbers. Then the dominant boundary condition can be written

$$(2.14) \quad \frac{\partial u}{\partial t} = \left(1 - \frac{t}{\sigma}\right) \left(1 - \frac{y}{\rho}\right)^{-1} \left[\sum_i G_i \frac{\partial u}{\partial x^i} + Hu \right],$$

where G_i ($i = 1, \dots, N-1$) and H are positive constants. Letting

$$(2.15) \quad \tau = -\sigma \log \left(1 - \frac{t}{\sigma}\right) = t + \frac{t^2}{2\sigma} + \dots,$$

we can write this

$$(2.16) \quad \frac{\partial u}{\partial \tau} = \left(1 - \frac{y}{\rho}\right)^{-1} \left[\sum_i G_i \frac{\partial u}{\partial x^i} + Hu \right], \quad x = 0.$$

Denote $\sum_i G_i$ by G .

In proving the convergence theorem we will actually assume that the left side of (2.16) dominates the right side. Since the series in (2.15) has positive coefficients, this will imply that the left side of (2.14) dominates the right

side, and hence that the boundary condition (2.9) will be dominated as required.

The dominating differential equation takes the form

$$(2.17) \quad \frac{\partial^2 u}{\partial x \partial t} = \left(1 - \frac{t}{\sigma}\right)^{-1} \left(1 - \frac{x+y}{\rho}\right)^{-1} \left[\sum_{i,k} A_{ik} \frac{\partial^2 u}{\partial x^i \partial x^k} + \sum_i B_i \frac{\partial u}{\partial x^i} + Cu + F \right],$$

where the A_{ik} , B_i , C and F are constants. Here only F depends on f_2 in (2.6) and we have therefore to find a radius of convergence independent of F . Let us assume that U is a function of τ , x , and y only. Then (2.17) becomes, with use of (2.15),

$$(2.18) \quad \frac{\partial^2 u}{\partial x \partial \tau} = \left(1 - \frac{x+y}{\rho}\right)^{-1} \left[A_{11} \frac{\partial^2 u}{\partial x^2} + A_{12} \frac{\partial^2 u}{\partial x \partial y} + A_{22} \frac{\partial^2 u}{\partial y^2} + B_1 \frac{\partial u}{\partial x} + B_2 \frac{\partial u}{\partial y} + Cu + F \right],$$

where the A 's are chosen anew if necessary. Since we can always increase them, we shall require that

$$(2.19) \quad A = A_{11} + A_{12} + A_{22} > G.$$

If we further assume that U is a function of the combination

$$w = \tau + \alpha(x+y)$$

alone, where $\alpha > 0$, then we can find an ordinary differential equation for $U(w)$ which dominates (2.18) and therefore still more dominates (2.6). To do this we replace $x+y$ in the denominator of (2.18) by $x+y+\tau/\alpha = w/\alpha$. Collecting terms in the ordinary differential equation to which (2.18) now leads, we find

$$(2.21) \quad \left(1 - \frac{w}{\alpha\rho} - \alpha A\right) U'' = BU' + \frac{C}{\alpha} U + \frac{F}{\alpha}.$$

Here primes denote differentiation with respect to w . We now choose α so small that

$$(2.22) \quad 1 - \alpha A > \frac{1}{2}.$$

According to (2.18) we would set B in (2.21) equal to $B_1 + B_2$. Since we can increase B freely without destroying the dominance over (2.6), we shall stipulate that

$$(2.23) \quad B > H \frac{1 - \alpha A}{1 - \alpha G} + \frac{1}{\alpha\rho}.$$

Defining $r = \alpha\rho(1 - \alpha A)$ we can write (2.21) in the form

$$(2.24) \quad U'' = \left(1 - \frac{w}{r}\right)^{-1} \left[\frac{B}{1 - \alpha A} U' + \frac{C}{\alpha(1 - \alpha A)} U + \frac{F}{\alpha(1 - \alpha A)} \right].$$

If in this equation $U(0)$ and $U'(0)$ are positive, then all coefficients of the series solution are positive.

Indeed, if we set

$$(2.25) \quad U = \sum_{n=0}^{\infty} a_n w^n,$$

the recursion formula for the a_n is seen to be

$$(n+2)(n+1) a_{n+2} = (n+1) \left[\frac{n}{r} + \frac{B}{1-\alpha A} \right] a_{n+1} + \frac{C}{\alpha(1-\alpha A)} a_n + \frac{F \delta_n^n}{\alpha(1-\alpha A)},$$

where the last term on the right is present only if $n = 0$. Assuming now that a_0 and a_1 are non-negative, we find

$$(2.26) \quad (n+2) a_{n+2} \geq \left[\frac{n}{r} + \frac{B}{1-\alpha A} \right] a_{n+1}.$$

This relation is used below.

Now consider the boundary condition (2.16). The formulae (2.9) and (2.12) show that the initial values for the u_n will be dominated if the left side of (2.16) dominates the right side. Still more will this hold if y in (2.16) is replaced by $x + y + r/\alpha = w/\alpha$. With this modification we get for $U(w)$ the condition

$$(2.27) \quad U'(w) \gg \left(1 - \frac{w}{\alpha p}\right)^{-1} [\alpha G U' + H U]$$

which will hold if

$$(2.28) \quad \left(1 - \frac{w}{\alpha p} - \alpha G\right) U'(w) \gg H U(w).$$

To verify that (2.28) implies (2.27) we recall that U is a series with positive coefficients; thus if we add to each side the series $\alpha G U'$ and then multiply on right and left by the series for $(1 - w/\alpha p)^{-1}$ we will not destroy the dominating relation.

To demonstrate (2.28) we calculate the coefficient of w^n on the left; it is

$$(1 - \alpha G)(n+1) a_{n+1} - \frac{1}{\alpha p} n a_n$$

which by (2.26), with $n+1$ changed into n , is not less than

$$(1 - \alpha G) \left[\frac{n-1}{r} + \frac{B}{1-\alpha A} \right] a_n - \frac{1}{\alpha p} n a_n$$

From (2.23) we find that this in turn exceeds

$$\begin{aligned} & \left[(1 - \alpha G) \left(\frac{n-1}{\alpha p(1-\alpha A)} + \frac{H}{1-\alpha A} \cdot \frac{1-\alpha A}{1-\alpha G} + \frac{1}{\alpha p(1-\alpha A)} \right) - \frac{n}{\alpha p} \right] a_n \\ & > \left[H + \frac{n}{\alpha p} \left(\frac{1-\alpha G}{1-\alpha A} - 1 \right) + \frac{\alpha G}{\alpha p(1-\alpha A)} \right] a_n. \end{aligned}$$

Since $G < A$ the middle term in the bracket is positive and thus the coefficient exceeds H which is the coefficient of w^n on the right of (2.28). This proves that (2.28) and (2.27) hold in general and hence that (2.16) and so (2.9) are dominated in the required way when $x = 0$.

For dominating power series we therefore choose a solution $U(w)$ having positive values for $U(0)$ and $U'(0)$. The radius of convergence of this series is equal to $r = \alpha\rho(1 - \alpha A)$, from the theory of linear differential equations, and this is independent of F .

Repeating this work at other points of C_0 we can show that the unique analytic solution thus found exists in a neighbourhood $t < \delta_1$, of C_0 . Here δ_1 is independent of the datum functions of the Cauchy problem as well as the mixed boundary condition. If we select any compact portion of T such that the above hypotheses are uniformly satisfied when any one of the characteristic surfaces G_t is chosen in place of G_0 then we can find a δ_1 which will serve for them all.

Combining the local solution just constructed with the solution of Cauchy's problem for the analytic case, we see that the resulting composite solution is analytic except possibly on G_0 . If the datum function $f(x')$ originally given satisfies the compatibility conditions of §1 up to the order k inclusive, then, by well-known properties of the discontinuities across characteristic surfaces of derivatives of u , it follows that u and its derivatives up to order k inclusive are continuous across G_0 . We state this result as a lemma:

LEMMA II. *Let the compatibility conditions up to order k inclusive be satisfied in the analytic case; then there exists a unique solution analytic for $0 < t < \delta_1$ except on G_0 , where the derivatives of order up to k inclusive are continuous.*

The domain of definition of this local solution will be extended in §4. We note that for the purposes of this local analytic solution it is sufficient to have (2.9) and thus v may be tangent to T .

3. Estimates of solutions. To extend the result to non-analytic equations and data, we give estimates of the square integrals of the solution and its derivatives up to a certain order. These are found by a modification of the method used by Krzyzanski and Schauder (5), which in turn is based on the work of Friedrichs and Lewy (3). For brevity we shall indicate only the alterations necessary for our purposes. In this section we take the geometric background to be Euclidean. It is also convenient to suppose that T is cylindrical in the sense that S spans T and the rim $C_0 = S \cap T$ is closed.

Since the Cauchy problem is regarded as solved, we can take as initial spacelike surface any spacelike surface which spans T , meeting T in the rim C_0 . We shall construct a family of surfaces S_t spanning T , with $S_0 \cap T = C_0$, and such that the given vector v is never tangent to the S_t . The direction field v is again assumed admissible; thus there exists a family G_t of characteristic surfaces, with $G_0 \cap T = C_0$, such that v is not tangent to the G_t . Let us extend

v to a field defined throughout a region of space containing the G_i ; if this region is sufficiently small we can, even in the analytic case, determine v analytically so that it is never tangent to G_i in the region. Denoting the minimum angle of v to G_i by θ_0 , we construct spacelike surfaces S_t as follows: S_t shall contain $C_t = T \cap G_i$; and S_t shall be inclined to G_i at an angle between $\frac{1}{3}\theta_0$ and $\frac{2}{3}\theta_0$ at every point. These surfaces S_t may be chosen to be analytic in the analytic case. Now we see that v is never tangent to the S_t .

We now set up a coordinate system on the family of surfaces S_t with equation $t = \text{const}$. We choose the coordinate network x^1, \dots, x^{N-1} on S_t in such a way that the parametric lines of t cross T at every point from inside to outside with increasing t . This can be achieved by a change of scale in a suitable "radial" coordinate in S_t , and does not alter the spacelike character of S_t . Since v is never tangent to S_t , we could take as N th coordinate, in place of t , a suitable parameter ξ^N along the integral curves of v . The transformation of coordinates so defined is clearly non-singular; and will be used below in certain surface integrals taken over T . By measuring arc along the v curves starting on T we ensure that T has in these coordinates the equation

$$T: \xi^N = 0.$$

However, this requires that v should not be tangent to T , which we therefore assume for the rest of this section. This condition has been anticipated in the form of the statement of Lemma I.

Let V_t be the region bounded by S , T , and S_t ; and let n_1, n_2, n_3 denote the outward Euclidean normals on the surface of S , T , and S_t respectively. If $\cos(nx^i)$ denotes the cosine of n with the parametric line of x^i , then we have

$$(3.1) \quad \cos(n_1 t) < 0, \quad \cos(n_2 t) > 0, \quad \cos(n_3 t) > 0.$$

We now multiply the differential equation (1.1) by $\partial u / \partial t$ and integrate over V_t . After some partial integrations we find

$$(3.2) \quad \int_{S+T+S_t} \left[2 \sum_{i,k=1}^N a^{ik} \frac{\partial u}{\partial x^i} \frac{\partial u}{\partial x^k} \cos(nx^k) - \sum_{i,k=1}^N a^{ik} \frac{\partial u}{\partial x^i} \frac{\partial u}{\partial x^k} \cos(nl) \right] dS \\ = \int_{V_t} \left(\Phi + f \frac{\partial u}{\partial t} \right) dV,$$

where Φ is a quadratic expression in u and its first derivatives, involving also the coefficients in (1.1) and their first derivatives.

A separate choice of variables is now made in each of the three surface integrals on the left in (3.2). The coordinates x^1, \dots, x^{N-1} are not changed but the last coordinate is taken to be

$$(3.3) \quad \eta_s^N = g_s(x^1, \dots, x^{N-1}, t), \quad s = 1, 2, 3.$$

where the g_s are the functions giving the equations of S , T and S_{t_0} when equated to zero. Thus

$$g_1 = t, \quad g_2 = \xi, \quad g_3 = t - t_0.$$

After some calculation, which we omit, it is possible to verify as in (5) that the integrand on the left in (3.2) becomes

$$\left[a(g_s) \left(\frac{\partial u}{\partial \eta_s} \right)^2 - \sum_{i,k=1}^{N-1} a^{ik} \frac{\partial u}{\partial x^i} \frac{\partial u}{\partial x^k} \right] \cos(n_s t)$$

where

$$a(g_s) = \sum_{i,k=1}^N a^{ik} \frac{\partial g_s}{\partial x^i} \frac{\partial g_s}{\partial x^k}$$

is the characteristic quadratic form. Since S and S_{t_0} are spacelike and T timelike, we see that

$$a(g_1) < 0, \quad a(g_2) > 0, \quad a(g_3) < 0,$$

in view of the convention of sign in (1.2). Noting that the quadratic form

$$\sum_{i,k=1}^{N-1} a^{ik} \xi_i \xi_k$$

is positive definite, and $\cos(n_s t)$ is positive, we see that for $s = 3$, that is for the integral over S_{t_0} , the integrand is negative definite. We shall now drop the subscript 0 on the t in S_t .

Since $\cos(n_s t)$ is positive, the term

$$- \sum_{i,k=1}^{N-1} a^{ik} \frac{\partial u}{\partial x^i} \frac{\partial u}{\partial x^k} \cos(n_s t)$$

is also negative definite. We transpose to the right side of (3.2) the other term in the integral over T and also the whole of the integral over S , and find, after changing the sign throughout,

$$\begin{aligned} & \int_{S_t} \left[-a(g_3) \left(\frac{\partial u}{\partial \eta_3} \right)^2 + \sum_{i,k=1}^{N-1} a^{ik} \frac{\partial u}{\partial x^i} \frac{\partial u}{\partial x^k} \right] \cos(n_s t) dS \\ (3.4) \quad & + \int_T \sum_{i,k=1}^{N-1} a^{ik} \frac{\partial u}{\partial x^i} \frac{\partial u}{\partial x^k} \cos(n_s t) dS \\ & = - \int_{V_t} \left(\Phi + f \frac{\partial u}{\partial t} \right) dV + \int_S + \int_T a(g_3) \left(\frac{\partial u}{\partial \eta_3} \right)^2 dS. \end{aligned}$$

The left hand side of this equation is now positive definite in all of the derivatives appearing, and in particular the integral over T on the left is non-negative. Thus we can drop this term provided \leq is substituted for the equality sign, and the inequality so obtained is in the right direction for our purposes.

Indeed, there is a positive constant c such that the left side of (3.4) exceeds

$$c \sum_{i=1}^N \int_{S_t} \left(\frac{\partial u}{\partial x^i} \right)^2 dS.$$

Now we find an upper bound for the right hand side by conventional methods, and so obtain an estimate

$$(3.5) \quad \sum_{t, k=1}^N \int_{S_t} \left(\frac{\partial u}{\partial x^i} \right)^2 dS < K \left[\int_{V_t} \sum_{i=1}^N \left(\frac{\partial u}{\partial x^i} \right)^2 dV + \int_{V_t} f^2 dV + \int_S \left[\sum_{i=1}^N \left(\frac{\partial u}{\partial x^i} \right)^2 + u^2 \right] dS + \int_T \left(\frac{\partial u}{\partial \xi^N} \right)^2 dS \right]$$

for some constant K independent of u . The expression Φ may contain u itself and so we have estimated integrals of the type

$$\int_V u^2 dV$$

by writing

$$u = u_0 + \int_0^t u_t dt'$$

and

$$\begin{aligned} \int_{V_t} u^2 dV &\leq 2 \int_{V_t} u_0^2 dV + 2 \int_{V_t} \left(\int_0^t u_t dt' \right)^2 dV \\ &\leq 2t \int_S u_0^2 dS + 2t^2 \int_{V_t} u_t^2 dV. \end{aligned}$$

These terms are incorporated on the right in (3.5). Integrating from $t = 0$ to $t = t_0$ in (3.5), we get

$$(3.6) \quad \int_{V_{t_0}} \sum_{i=1}^N \left(\frac{\partial u}{\partial x^i} \right)^2 dV < K_1 \left[\int_{V_{t_0}} f^2 dV + \int_{S_0} \left[\sum_{i=1}^N \left(\frac{\partial u}{\partial x^i} \right)^2 + u^2 \right] dS + \int_T \left(\frac{\partial u}{\partial \xi^N} \right)^2 dS \right],$$

provided that t_0 is sufficiently small. Replacing the first term on the right in (3.5) by means of (3.6), we finally get a similar estimate for the surface integral on the left in (3.5).

Similar estimates for the higher derivatives of u are needed; as they can be found by modifying the calculations of (5, §3) in the manner indicated above, the details will be omitted. We quote the result as follows: Let $D_h u$ denote a typical partial derivative of order h , and let ρ be a positive integer. Let the coefficients a^{ik} , b^k , c , and f of (1.0) have bounded derivatives up to and including the order $N+1$; and let these derivatives of order up to $N+\rho-1$ be square integrable over the domain. Then there holds the estimate

$$(3.7) \quad \sum_{h=0}^{N+\rho} \int_V (D_h u)^2 dV < K_\rho \left[\sum_{h=0}^{N+\rho} \int_V (D_h f)^2 dS + \sum_{h=0}^{N+\rho-1} \int_T \left(D_h \frac{\partial u}{\partial \xi^N} \right)^2 dS + \sum_{h=0}^{N+\rho} \int_S (D_h u)^2 dS \right],$$

and a similar estimate holds for

$$(3.8) \quad \sum_{h=0}^{N+p} \int_{S_t} (D_h u)^2 dS_t.$$

In these formulae the summation \sum over h is to be taken over all derivatives of the order indicated. However, in the integral over T the summation \sum' over h shall include only derivatives tangential to T and so only one differentiation with respect to $\xi^N = \eta^{N_2}$ will appear. Again, by integrating (3.8) over t , we find the sharp form

$$(3.9) \quad \sum_{h=0}^{N+p} \int_{V_t} (D_h u)^2 dV < t\tilde{K} \quad]$$

of the estimate for the volume integrals. This is to be used in connection with quasi-linear equations, which we shall mention in §5.

4. Extension of the domain. The following lemmas of Schauder and Krzyzanski (5, §6) will be used here and in the next section. Let R_k denote the class of functions $v(x^i)$ having absolutely continuous derivatives of order $\leq k-1$ in the sense of Tonelli; and quadratically integrable derivatives of order $\leq k$; all on a given closed domain such as the region V_t . Such a function is absolutely continuous in the above sense when it is absolutely continuous on almost every parametric line of the chosen coordinates.

With the norm

$$(4.1) \quad ||v||_k = \sum_{h=0}^k \int_{V_t} (D_h v)^2 dV,$$

the class R_k becomes a linear space R_k . We have

LEMMA III. *Polynomials $p(x^i)$ are dense in R_k . Provided that $v \in C^r$ on a compact subset V_1 of V , an approximating sequence $p_j(x^i)$ can be found such that the derivatives of order $\leq r$ of the p_j converge uniformly to the derivatives of v .*

LEMMA IV. *Let $k \geq N+1$, and let $v_n \in R_k$ be a sequence with uniformly bounded norms $||v_n||_k < K$. Then there exists a subsequence v_{n_j} uniformly convergent to a limit $v \in R_k$.*

The uniform convergence is established by Lemma I of (9) while the fact that v belongs to R_k follows (1) from the theory of strong convergence in L^2 .

We now extend the domain of definition of the solution of Lemma II. Let $0 < \delta < \delta_1$ and let us divide the larger domain V into p slices V_j of width δ :

$$V_j: (j-1)\delta \leq t \leq j\delta, \quad j = 1, 2, \dots, p.$$

Let S_j be the surface $t = j\delta$ and $C_j = S_j \cap T$. In each of these domains we shall construct solutions which will subsequently be pieced together. The solution in the large so found will be class R_k , provided that the compatibility conditions of order $\leq k$ hold initially, and that $k \geq N+2$.

Let u_1 be the local solution defined in V_1 . We cannot apply Lemma II to V_2 by taking values of u_1 and $\partial u_1 / \partial t$ on S_2 as Cauchy data since these may

not be analytic in the large. However, as is remarked in (5, §7), u_1 and $\partial u_1/\partial t$ have derivatives of all orders on S_1 except on $G_0 \cap S_1$; and satisfy the compatibility conditions of all orders on C_1 . We therefore approximate them by polynomial sequences ϕ_{1s} and ψ_{1s} such that (a) the derivatives of orders $\leq k$ and $k-1$ respectively converge uniformly on S_1 while (b) on C_1 the derivatives of orders $\leq k+N(p-2)$ converge. These approximations are possible, by Lemma III.

To each pair ϕ_{1s}, ψ_{1s} there corresponds a solution u_{2s} of the Cauchy problem with data on S_1 . We define these solutions throughout V_2 by selecting analytic boundary values χ_s which satisfy the compatibility conditions relative to u_{2s} on C_1 , up to the order $k+N(p-2)$. Thus the derivatives of χ_s of order $\leq k+N(p-2)$ converge to the corresponding derivatives of f on C_1 . Now Lemma II shows that u_{2s} is defined and of class $R_{k+N(p-2)}$ in V_2 .

To extend these solutions to V_3 and beyond, we approximate u_{2s} and $\partial u_{2s}/\partial t$ on S_2 by sequences ϕ_{3sr} and ψ_{3sr} of polynomials. By Lemma III these approximations can be made uniform for derivatives of order $\leq k+N(p-2)$. Again we define solutions u_{3sr} in V_3 , with Cauchy data ϕ_{3sr} and ψ_{3sr} , and boundary data χ_{3sr} , where χ_{3sr} is a polynomial satisfying the appropriate compatibility conditions of order $\leq k+N(p-2)$. By Lemma II, the solutions u_{3sr} exist in V_3 and are of class $R_{k+N(p-2)}$ there. Also, by (3.7),

$$(4.2) \quad \|u_{3sr}\|_{k+N(p-2)} \leq K,$$

where K is independent of r and s . By Lemma IV, there exists for each s a subsequence u_{3sr_p} convergent to a limit u_{3s} of class $R_{k+N(p-2)}$ in V_3 ; thus u_{3s} also satisfies the differential equation and the estimate (4.2). We now approximate to values of u_{3s} and $\partial u_{3s}/\partial t$ on S_3 in order to define Cauchy data for a sequence of solutions u_{4sr} of class $R_{k+N(p-3)}$ in V_4 . The approximations to the boundary condition are again of class $C_{k+N(p-3)}$ and there exists a sequence of solutions u_{4s} of class $C_{k+N(p-4)}$ in V_4 which satisfy a uniform estimate of the type (4.2).

Proceeding in this way we define a sequence of solutions u_{js} of class $R_{k+N(p-j)}$ in V_j , all satisfying an estimate of the type (4.2). We now piece together the solutions u_{js} , for fixed s , to give solutions u_s defined in $V_2 + V_3 + \dots + V_p$, which are of class R_k . By Lemma IV there exists a subsequence uniformly convergent to a limit u of class R_k in $V_2 + V_3 + \dots + V_p$. Assuming that $k \geq N+2$, this function u has continuous first and second derivatives, and satisfies the differential equation and the boundary condition. Also, by the manner of its construction, this solution merges with the original solution in V_1 to yield a solution U of class R_k in $V = V_1 + V_2 + \dots + V_p$. This completes the proof that the solution can be extended to a domain of arbitrary extent.

5. The non-analytic case. In the differential equation (1.1) let all coefficients and f be of class R_{k-1} (where $k \geq N+2$) in V_1 and let the data

of the mixed problem be of class C^{k+N} . Let W (in Schauder's notation) be a domain which contains S and is contained in the region of dependence on S alone. We suppose that the differential equation is of class C^{k+N} in W . Finally, the conditions of compatibility up to order k inclusive shall be satisfied. We state the result in this case as follows.

THEOREM. *There exists a solution u of the given mixed problem, which is of class R_k in V , of class R_{k+N} and C^k in W , and which satisfies an estimate of the form (3.7).*

The proof involves approximation to the coefficients of the differential equation by polynomials. From Lemmas III and IV we see that the approximating coefficients a_s^{ik} , b_s^i , c_s and f_s ($s = 1, 2, \dots$), together with derivatives up to order $k - 1$ inclusive, can be chosen to converge to their respective limits (a) in V , in mean and (b) in W , uniformly. As in the preceding extension of the domain of §4, the data can be approximated by polynomials which retain the k compatibility conditions. According to §4, there exists a solution u_s of the approximate problem, with

$$\|u_s\|_k < K.$$

From Lemma III we infer the existence of a uniformly convergent subsequence tending to a limit $u \in R_k$. Since $k \geq N + 2$, u is C^2 and satisfies the differential equation and the boundary conditions. In fact u is of class R_{k+N} in W as follows from the theory of the Cauchy problem (9). That the solution is unique follows from the estimate (3.7).

A boundary condition of the third kind (in potential theory) can be reduced to that treated here. If

$$\frac{\partial u}{\partial \nu} + hu = f,$$

where h and f are functions of position on T , then the reduction in (2) will apply.

With boundary value problems for hyperbolic equations there is an evident analogy with potential theory, and Hadamard (4, p. 248 ff) discusses these problems in that light. However, the case of a plane boundary treated by him is essentially easier than the general case since in effect it can be solved by the method of images. The result found in this paper has a greater generality than one would expect by this analogy, since the case of the oblique derivative, which in potential theory requires special methods, is included.

In conclusion we note that Schauder has also treated the quasi-linear and non-linear mixed problems with the values of u assigned on T , (5, 10). His methods extend without difficulty to the boundary condition studied here. Indeed, in the quasi-linear case, the linear solution is used to define a functional transformation, and then with the help of the sharp estimate (3.9) it is shown that a fixed point of the transformation, and hence a solution, exists for sufficiently small domains. The non-linear problem is reduced by differentiation to a quasi-linear integro-differential system which can be solved under the

same conditions as the quasi-linear hyperbolic equation. Schauder's proof of the integrability conditions for this system, which establish the existence of the solution for the non-linear equation, requires no modification in the present case.

6. Removal of the compatibility conditions. The preceding result has been stated under conditions similar to those in (5) with the boundary condition of the first kind. In both of these theorems the compatibility conditions of order up to k are somewhat inappropriate in view of the theory of discontinuities of derivatives of solutions of hyperbolic equations. To remove this limitation, we shall need to strengthen the differentiability conditions.

Consider, therefore, the first boundary condition when $q + 1$ ($0 \leq q < k$) compatibility conditions hold. We shall actually treat the case $q = 0$ since the continuity across G_0 of the derivatives up to order q is easily established later in the appropriate cases. Thus, taking the Cauchy data to vanish and considering the homogeneous differential equation, we assume only that $f(x^i)$ vanishes to the first order on C_0 . Let the differential equation and data be of class C^{2k} , and let us reduce this problem to that treated in (5) by setting

$$u = u_1 + v.$$

We shall arrange that $L(v)$ be C^k everywhere and that u shall satisfy a boundary condition on T which is compatible of order k .

More precisely, we set $v \equiv 0$ in the Cauchy domain between S and G , and require that $L(v)$ should vanish to the order k as G is approached from above. The function v itself shall be continuous, shall vanish on G and shall satisfy on T the boundary condition

$$v = \sum_{n=1}^k f_n t^n,$$

where the f_n are the coefficients of f in a Taylor series expansion in powers of t . For u_1 we now have

$$u_1 = \tilde{f} \equiv f - \sum_{n=1}^k f_n t^n \quad \text{on } T.$$

Since we have assumed $L(u) = 0$ in the reduced form of the boundary value problem, the first compatibility conditions for u_1 will be the vanishing of the appropriate derivatives of f . Thus the problem for u_1 is of the above type, since we have in effect taken $u = 0$ in the region W of the theorem. This shows that the problem is in this case reduced to finding v .

For this purpose we note that all functions and coefficients can be expanded in a Taylor series of powers of t (where $t = 0$ is the equation of G) up to terms of order t^k and with a remainder of this order in t . The coefficients of terms containing t^r are derivatives of order r , and so are C^{2k-r} . Hence all such coefficients are C^k . Following Hadamard (4, pp. 78-79), we construct the first k terms of the series in powers of t , of the problem

$$L(w) = 0,$$

with $w = 0$ on G and $w = f$ on T .

By the manner of its construction, the function

$$v = \sum_{n=1}^k w_n t^n$$

satisfies

$$v = \sum_{n=1}^k f_n t^n \quad \text{on } F,$$

and

$$L(v) = t^{k+1}r$$

where r is a C^k remainder term, in the region between G and T . Thus $L(v)$ has continuous derivatives up to order k across G and the reduction is established.

If the first $q + 1$ compatibility conditions hold, then it is easy to show that u is C^q across G , $q < k$, by considering the discontinuities of successive derivatives across G and noting that since they vanish on C they must vanish along all bicharacteristics issuing from C . We may now state the existence theorem of (5) with this modification.

Let the differential equation and boundary datum function f be C^{2k} , $k \geq N + 2$, and let the first $q + 1$ compatibility conditions hold. Then there exists a unique solution of $L(u) = 0$ in V with given Cauchy data and with $u = f$ on T . The solution is of class R_k in V except that if $q < k - N$ it is of class C^q across G .

The domain of this solution is, however, restricted to the domain wherein v has been defined and so does not include any multiple points of the characteristic surface G .

A similar reduction for the boundary condition of the second kind, considered in this paper, is possible. Here, however, it is not necessary that any compatibility condition should hold. We calculate v as the first k terms of the analytic series expansion in §2, and proceed as above. The result may be stated as follows when q compatibility conditions hold.

Let the differential equation be C^{2k} and the boundary datum function be C^{2k-1} , $k \geq N + 2$. Let the first q compatibility conditions hold on C . Then there exists a unique solution of $L(u) = 0$ in V with given Cauchy data and with

$$\frac{\partial u}{\partial v} = f$$

on T . The solution is of class R_k in V , except that if $q < k - N$ it is of class C^q across G .

Again the domain is limited by multiple points or self-intersections of the characteristic surface G .

REFERENCES

1. S. Banach and S. Saks, *Sur la convergence forte dans les champs L^p* , *Studia Math.*, 2 (1930), 51-57.
2. G. F. D. Duff, *Uniqueness in boundary value problems for the second order hyperbolic equation*, *Can. J. Math.*, 8 (1956), 86-96.
3. K. Friedrichs and H. Lewy, *Ueber die Eindeutigkeit und das Abhängigkeitsgebiet der Lösungen beim Anfangswertproblem linear hyperbolischer Differentialgleichungen*, *Math. Ann.*, 98 (1928), 192-204.
4. J. Hadamard, *Lectures on Cauchy's problem in linear partial differential equations* (New York, 1952).
5. M. Krzyzanski and J. Schauder, *Quasilineare Differentialgleichungen zweiter Ordnung vom hyperbolischen Typus: Gemischte Randwertaufgaben*, *Studia Math.*, 6 (1936), 162-189.
6. O. Ladyshenskaya, *Smesannaya zadaca dlya gyperboliceskova uravneniya* (Moscow, 1953).
7. J. Leray, *Hyperbolic Differential Equations* (Princeton, 1953).
8. A. Robinson and L. Campbell, *Mixed problems for hyperbolic partial differential equations*, *Proc. London Math. Soc.* (3), 5 (1955), 129-147.
9. J. Schauder, *Das Anfangswertproblem einer quasilinearen hyperbolischen Differentialgleichung . . .*, *Fund. Math.*, 24 (1935), 213-246.
10. ———, *Gemischte Randwertaufgaben bei Partiellen Differentialgleichungen vom hyperbolischen Typus*, *Studia Math.*, 6 (1936), 190-198.

University of Toronto

NONPARAMETRIC METHODS in STATISTICS

By D. A. S. FRASER, *University of Toronto*. Unifies and explains nonparametric methods from diverse areas of statistical activity. In the early chapters the author develops the statistical theories of hypothesis testing, confidence regions, and tolerance regions. Full coverage is given to limiting distributions, and a complete chapter details limiting methods for nonparametric test statistics. One of the *Wiley Publications in Statistics*, Walter A. Shewhart and S. S. Wilks, Editors. 1957. 299 pages. \$8.50.

IRRATIONAL NUMBERS

By IVAN NIVEN, *University of Oregon*. Offers an exposition of some central results on irrational numbers, rather than an exhaustive treatment of the subject. The main emphasis is on those aspects of irrational numbers commonly associated with number theory and Diophantine approximation. Number 11 in the series of the *Carus Mathematical Monographs*. 1956. 167 pages. \$3.00.

PRINCIPLES and TECHNIQUES of APPLIED MATHEMATICS

By BERNARD FRIEDMAN, *New York University*. Takes a significant step in bridging the gap between pure and applied mathematics. It (1) demonstrates how the abstract theory of linear operators can be used to unify the techniques of applied mathematics; and (2) explains specific techniques which can be used to obtain explicit solutions of partial differential equations. A publication in the *Wiley Applied Mathematics Series*, I. S. Sokolnikoff, Editor. 1956. 315 pages. \$8.00.

DIFFERENTIAL EQUATIONS, 3rd Edition

By HARRY W. REDDICK and DONALD E. KIBBEY, both of *Syracuse University*. More than half rewritten, the third edition has new features geared to modern teaching requirements. Among these are a chapter on partial differential equations, a new section on the adjoint equation, and many fresh problems integrated with the discussions. Explanations are clear and rigorous; problems illustrate principles and stress applications. 1956. 304 pages. \$4.50.

STATISTICAL ANALYSIS of STATIONARY TIME SERIES

By ULF GRENANDER, *University of Stockholm*; and MURRAY ROSENBLATT, *Indiana University*. Fulfills two major purposes: (1) it introduces theoretical statisticians to an approach to time series analysis that is essentially different from most of the techniques used by analysts in the past; and (2) it presents a unified treatment of modern methods that are being used increasingly in the physical sciences and technology. 1957. 300 pages. \$11.00.

University of Toronto Press, Toronto, Ontario

In Canada:

Renouf Publishing Company, Montreal, Quebec

Recent

OXFORD PUBLICATIONS

DIFFERENTIAL CALCULUS. By W. L. Ferrar. 306 pp.,
21 text-figures. \$4.25

The book is divided into three sections, on the foundations of the differential calculus, on functions of one variable, and on functions of two or more variables.

**ORDINARY NON-LINEAR DIFFERENTIAL EQUATIONS
IN ENGINEERING AND PHYSICAL SCIENCES.** By N. W.
McLachlan. *Second Edition.* 282 pp., 122 text-figures. \$5.25

A new and enlarged edition of a popular text-book that includes additional material on stability criteria.

AN INTRODUCTION TO LINEAR ALGEBRA. By L. Mirsky.
445 pp. \$5.25

Applications to geometry are stressed throughout the book; and such topics as rotation, reduction of quadrics to principal axes, and classification of quadrics are treated in some detail.

THE MATHEMATICS OF DIFFUSION. By J. Crank.
355 pp., 120 text-figures. \$7.50

This book is written for those who are interested in the mathematical work which has been done in diffusion within the last ten years.

MATHEMATICAL ANALYSIS. By D. A. Quadling. 272 pp.,
51 text-figures. \$3.75

The book makes the transition from methods of elementary calculus to those applicable to university courses on mathematical analysis.

*A complete list of Oxford Books on Mathematics
is available on application*

OXFORD UNIVERSITY PRESS

480 University Avenue

Toronto 2

